MOLECULAR CHARACTERISATION OF THE PNEUMONIA VIRUS OF MICE GLYCOPROTEIN GENES.

by J. S. RANDHAWA.

A THESIS SUBMITTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

OF THE UNIVERSITY OF WARWICK.

RESEARCH CONDUCTED IN THE

DEPARTMENT OF BIOLOGICAL SCIENCES.

September, 1993

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank Dr. Andrew Easton for supervision of this project and for his proof-reading of this thesis and for the financial support of the Medical Research Council. I would also like to acknowledge the support of my parents, brothers and wife during the course of this PhD. Finally, I wish I could thank my lab-mates, who shall remain nameless, for their generous advice and at times light entertainment, but then I would be lying. To all who read this thesis, live long and prosper.

DECLARATION

I hereby declare that all of the results presented within this thesis were obtained by myself under the supervision of Dr. Andrew Easton, with the exception of those instances where the contribution of others has been acknowledged. These results have not been submitted for a degree at any other institution.

ABBREVIATIONS

ATP : Adenosine triphosphate.

BrdU : 5-Bromo-2'-deoxyuridine

BSA : Bovine serum albumin.

CTP : Cytidine triphosphate.

DTT : Dithiothreitol.

dATP : 2'-deoxyadenosine 5'-triphosphate.

dCTP : 2'-deoxycytidine 5'-triphosphate.

dGTP : 2'-deoxyguanosine 5'-triphosphate.

dTTP : 2'-deoxythymidine 5'-triphosphate.

FCS : Bovine foetal calf serum.

GMEM : Glasgow modified Eagle's medium.

GTP : Guanosine triphosphate.

HEPES : N-2-hydroxyethylpiperazine-N-2-ethane sulphonic acid.

IPTG : Isopropyl-thio-ß-D-galactoside.

PBS : Phosphate buffered saline.

pfu : Plaque forming units.

PMSF : Phenylmethylsulphonylfloride.

TEMED : N, N, N', N'-tetra-methylethylenediamine.

TTP : Thymidine triphosphate.

X-Gal : 5-Bromo-4-chloro-3-indolyl-ß-D-galactoside.

SUMMARY

The molecular characterisation of the major glycoproteins of pneumonia virus of mice (PVM) was undertaken to identify the molecular basis for different pathogenicities for two strains of PVM. One is highly pathogenic and passaged entirely in mice (strain J3666) whilst the other is non-pathogenic and has been passaged entirely in tissue culture (strain 15).

Characterisation of the fusion proteins of these two strains revealed little amino acid changes that may account for their pathogenicities. Of the 537 amino acid long fusion protein, only four amino acid changes were observed between the two strains. None of the changes were located within the region encoding the cleavage site of the F_0 polypeptide and the altered amino acids were distributed throughout the fusion protein. The amino acids changes were considered to be fairly conservative and may represent adaptation of the two strains to their different modes of passaging. However, the importance of these changes with regards to cleavage cannot be ruled out since one or more of these changes may be located within the 3-dimensional structure of the F_0 cleavage site.

Nucleotide sequence analysis of the attachment (G) glycoproteins of these two strains of PVM showed only two amino acids changes within the extracellular domain of the protein. However, the G proteins of the two strain differed from each other in that the G protein of strain J3666 possessed a N-terminal extension compared to that of strain 15. This amino acid extension forms the cytoplasmic domain of the G protein in strain J3666 but appears to be lacking in that of strain 15. Furthermore, the G proteins of both strains of PVM were synthesised from the second, but not the same, initiation codon. In PVM strain J3666, the first available initiation codon coded for a short polypeptide of 12 amino acids whereas that of strain 15 coded for a short polypeptide of 33 amino acids.

Expression of the G genes *in vitro* produced polypeptides which were approximately 3K larger than their predicted molecular weights. Synthesis of polypeptides due to internal initiation during translation were observed for the G gene of strain J3666. Expression of the G genes using recombinant vaccinia virus expressing the T7 RNA polymerase showed that the G proteins of both strains of PVM were expressed on the surface of cells transfected *in vitro*. Thus, the signals required for the correct processing and transport of the G proteins of PVM are not contained within the cytoplasmic domain of the proteins.

The first steps of an investigation of possible ribosomal frameshifting event occurring within the G gene of PVM strain 15 was undertaken. The initiation codon of the major and short ORF's of the G gene of PVM strain 15 were mutated in a variety of combinations with a view to fusing these mutated fragments to the 5' end of the *E. coli* β-galactosidase gene that lacks its own initiation codon.

CHAPTER 1

INTRODUCTION

1.1. GENERAL :

The *Paramyxoviridae* family of viruses are classified in group V of the Baltimore classification of viruses (Baltimore, 1971, Kingsbury, 1978, Francki *et al*, 1991). These viruses have a negative sense single stranded RNA genome that is transcribed from a single polymerase binding site at the 3' end of the RNA molecule. Early work led these viruses to be classified into three genera (paramyxovirus, morbillivirus and pneumovirus) based on the characteristics of their attachment glycoprotein. The names paramyxo-, morbilli- and pneumo- are derived from the Greek: para meaning akin or closely related, myxo meaning 'mucus', morbilli from the plural of Latin morbillus - diminutive of morbus, and pneumo meaning 'breath'. The paramyxoviruses were demonstrated to have both haemagglutination and neuraminidase activity. Mumps virus, parainfluenza viruses 1, 2 and 3, Sendai virus and Newcastle disease virus are the best known members of this genus. Measles virus, canine distemper virus, rinderpest virus and peste de petits ruminants virus are classified as morbilliviruses and have haemagglutination but not neuraminidase activity. The prototype pneumovirus, human respiratory syncytial virus (HRSV), was initially classified into a separate genus since it did not demonstrate haemagglutination or neuraminidase activity. Subsequent analysis of the viral polypeptides supported this classification although another pneumovirus, pneumonia virus of mice (PVM) exhibited haemagglutination activity. Pneumoviruses appear as lipid-enveloped spherical, pleomorphic or filamentous particles and are associated with infections of the respiratory tract. HRSV is a major cause of hospitalisation due to respiratory infection in infants under the age of two. It is for this reason that HRSV has been studied extensively. The morphology of PVM (Compans *et al*, 1967; Berthiaume *et al*, 1974) and the number of viral encoded polypeptides (Cash *et al*, 1977, 1979; Ling and Pringle, 1989a) resemble those of HRSV. Other members of this genus are turkey rhinotracheitis virus (TRTV) and bovine respiratory syncytial virus (BRSV). Analysis

of the virus polypeptides and sequence data available for the genes of paramyxoviruses, morbilliviruses and pneumoviruses have supported classification of the *Paramyxoviridae* family of viruses into three genera.

1.2. CHARACTERISTICS OF PVM :

PVM was first isolated in 1939 from albino Swiss mice. The virus was shown to be extremely labile in that incubation at 56°C for 30 minutes completely inactivated infectivity whereas incubation at room temperature reduced the infectious titre by a hundred-fold. However virus could be stabilised with the addition of 10% horse serum although incubation at room temperature reduced the infectious titre by five-fold (Horsfall and Hahn, 1939, 1940). A decrease in virulence of PVM was observed following inoculation of mice with tissue culture passaged virus (Horsfall and Hahn, 1940). Haemagglutination and haemadsorption activities have been found with PVM (Mills and Dochez, 1944, Compans *et al*, 1967).

1.3. PREVALENCE AND INFECTION OF PVM :

PVM was inadvertently isolated from control mice during attempts to isolate viruses from patients suffering from atypical pneumonia following intranasal inoculation of mice. The isolated virus was shown to cause an inapparent or persistent infection in mice. However, following passage in lung tissue, intranasally inoculated healthy recipient mice developed a fatal pneumonia (Horsfall and Hahn, 1939, 1940). Neutralising antibodies against PVM have been found in cotton rats, rabbits (Eaton and van Herick, 1944), guinea pigs, hamsters, chimpanzees and monkeys (Horsfall and Curnen, 1946) although no apparent disease was found in any of these animals. Horsfall and Curnen (1946) demonstrated that PVM can cause pneumonia following intranasal inoculation of cotton rats and hamsters that did not exhibit neutralising antibodies against PVM. In a study of different viruses prevalent in laboratory animal colonies for the years 1978-1979, Gannon and Carthew (1980) showed PVM to be

present in 61% of rats and 59% of mice examined. No anti-PVM antibodies were found in any of the guinea pigs tested. Evidence for neutralising antibodies against PVM in 27-35% of human sera was reported (Horsfall and Hahn, 1940, Horsfall *et al*, 1943, Horsfall and Curnen, 1946) although Pringle and Eglin (1986) reported neutralising antibodies to PVM in 75% of adult human sera. Similar levels of seroconversion were observed with sera from Nigeria (Pringle and Eglin, 1986). The proportion of seropositive children increased with age as seen for HRSV. The results suggest exposure to PVM, or an antigenically related virus, is common. Levels of neutralising antibodies were higher in patients suffering from Paget's disease of the bone however the clinical consequence of PVM infection in humans is unknown. Although PVM is serologically distinct from HRSV, antigenic cross-reactivity has been found between the N (Gimenez *et al*, 1984, Ling and Pringle, 1989a) and P proteins (Ling and Pringle, 1989a) of these two viruses.

Weir *et al* (1988) observed a wasting disease accompanied by pneumonia in intranasally inoculated mice. Necropsy found a decrease in fat deposits and muscle mass, diffuse pulmonary oedema (swelling of part of the body due to congestion of excess body fluid), splenomegaly (abnormal enlargement of the spleen), cyanosis (blue coloration of the skin due to lack of oxygen in the blood supply), dyspnea (breathlessness or shortness of breath after slight physical effort due to diminished capacity of the lungs to cope with the demands put on them) and tachypnea (rapid and shallow breathing) (Horsfall and Hahn, 1939, 1940, Weir *et al*, 1988). Changes in the lungs were localised to the lobar regions where thickening and enlargement of the alveolar walls, intra-alveolar haemorrhage, fibrin deposition, accumulation of macrophages and large mononuclear cells with the alveolar spaces were observed (Carthew and Sparrow, 1980, Weir *et al*, 1988). Tissue staining with anti-PVM antibodies revealed the presence of PVM in both the bronchial epithelium and the alveolar cells and persistence of the virus only in the alveolar wall (Carthew

and Sparrow, 1980). However, Weir *et al* (1988) observed the conducting airways to be normal with PVM replication localised solely within the alveolar cells. This data supports the findings of Horsfall and Hahn (1940) and Horsfall and Ginsberg (1951) indicating replication is restricted to the lungs. Smith *et al* (1984) failed to observe pneumonia in PVM infected mice and determined replication to occur in the upper respiratory tract. Differences in these data can be reconciled to the use of different strains of mice, different virus strains, age of mice and duration of infection. Mice used in studies by Smith *et al* (1984) were 25 days old and inoculated with PVM by placing the virus on the external nares whereas those used by Weir *et al* (1988) were over 2 months of age and inoculated intranasally.

1.4. MORPHOLOGY OF PVM :

Electron microscopy studies on PVM infected cells revealed three forms of the virus. The most common were the filamentous forms with a uniform diameter of 100-200 nm with a total length of 2-3 μm (Compans *et al*, 1967, Berthiaume *et al*, 1974). Weir *et al*, 1988 calculated these filamentous forms to be 85-90 nm in diameter and up to 3 μm in length. The filaments protruded from the cell surface with the distal end having an enlarged bulbous tip. Although filamentous forms of HRSV were observed, no enlargement of either end was observed (Berthiaume *et al*, 1974, Weir *et al*, 1988). The significance of the bulbous structure at one end of the virus is unknown. The circular profiles of the virus were 80-120 nm in diameter and were cross-sections of spherical particles and not cross-sections of the filamentous forms (Compans *et al*, 1967). The viral envelope is covered in spike-like projections 12 nm in length, 10 nm apart in HRSV but only 6 nm apart in PVM (Compans *et al*, 1967, Berthiaume *et al*, 1974, Weir *et al*, 1988). The third type of structure observed was a modification of the filamentous form in which PVM appeared as a horseshoe shape with two narrow filamentous structures at either end (Berthiaume *et al*, 1974).

The internal structures appeared as electron dense strands extending across the circular form and as coiled strands within the filamentous forms which ran parallel to the long axis (Compans *et al*, 1967, Berthiaume *et al*, 1974, Weir *et al*, 1988). Turns of the helical structure were well separated indicating a single-stranded genome. The nucleocapsids of HRSV and PVM were approximately 13.5 nm in diameter with a pitch of the helix of 6.5 nm (Berthiaume *et al*, 1974) which differentiated them from other paramyxoviruses. Infected cells contained dense cytoplasmic inclusions with a thread-like appearance similar to those seen under the light microscope when stained with acridine orange. Additionally, vesicles were found at the surface of the infected cells between the regions of virus budding (Compans *et al*, 1967). This raises the interesting possibility that they represent either defective interfering particles with the nucleocapsid as a minor component or that they are regions where the matrix protein interacts with a cell membrane component thus forming vesicles analogous to the interaction of the matrix (M) protein of VSV with cell membranes in transfected cells which results in evagination (Blondel *et al*, 1990).

1.5. PARAMYXOVIRUS REPLICATION/TRANSCRIPTION :

Comparison of the terminal 3' and 5' sequence of HRSV with those of other paramyxoviruses shows some degree of homology. A feature common to all leader sequences of paramyxoviruses is the richness in U sequences (Fig. 1.1 ; Mink *et al*, 1991). Whereas the first two nucleotides of the terminal 3' end are identical to those for some paramyxoviruses (measles virus, Sendai virus, PIV3, BPIV3, NDV and PIV2), the first three nucleotides (UGC) are identical to those of VSV and rabies virus (RaV). Another feature of paramyxoviruses is the presence of sequences in the 5' trailer complimentary to regions in the leader sequence. Thus, the possibility exists for circular forms of the viral genome. Indeed, Udem and Cook (1984) have shown the presence of such a form of the measles virus genome in infected cells using electron microscopy.

```
3' UGCGCUUUUUUUACGCAUGUUGUUUGAACGUAUUUGGUUUUUUUUAC
   ||| ||||||| | || ||| |||||   ||      ||   | |
5' ACGAGAAAAAAAGUGU-CAA-AAACUAAUAUCUCGUAAUUUAGUU
```

Fig. 1.1 : HRSV leader and trailer nucleotide sequences.

Complementarity between 3' and 5' end sequences of HRSV. Nucleotide sequence is shown in genomic RNA sense. The vertical lines represent pairing between 3' and 5' end nucleotides (Taken from Mink *et al*, 1991).

The discovery of leader RNA's and the presence of a nucleation site for encapsidation by the nucleocapsid protein (N) protein at the 5' end of the leader RNA in VSV led to a theory for the mechanism of viral transcription in negative stranded, non-segmented viruses (Banerjee, 1987, Blumberg *et al*, 1991) in which the N protein acts as an anti-terminator of the RNA polymerase. Synthesis of the leader RNA was followed by encapsidation with the N protein while absence of sufficient N protein resulted in the termination of leader RNA elongation after which synthesis of viral mRNA was initiated. Thus factors affecting the concentration of cytoplasmic N protein would have a significant influence on the replicative/transcriptional procedure. It has been demonstrated that the N protein is capable of forming aggregates thereby reducing the concentration of free N protein. Interaction of the NS protein (the rhabdovirus equivalent of the *Paramyxovirus* phosphoprotein protein) with free N protein may prevent formation of aggregates and thus keep it in a form capable of binding to RNA (Banerjee, 1987). Howard and Wertz (1989) demonstrated the effect of the N protein concentration of VSV in replication of DI nucleocapsid *in vitro*. At low concentrations of the N protein, DI replication was supported whereas at higher concentrations aggregates of the N protein were observed accompanied by an

inhibition of DI replication. Co-synthesis of the NS and N proteins prevented aggregation of the N protein and supported replication of the DI genome even when concentrations of the N protein were high. The results demonstrate the importance of controlling the concentrations of the N protein during virus replication. Additionally the concentration of ATP was found to play a role in regulation of viral transcription. The absence of detectable leader RNA's in most paramyxoviruses has raised doubts as to the general application of this theory for all negative stranded, non-segmented viruses.

Blumberg *et al* (1991) postulate the presence of two promoter sites within the leader sequence, one responsible for initiation of replication with the other responsible for initiation of transcription. Additionally, by analogy to tobacco mosaic virus (TMV) which has a nucleation site located approximately 900 nucleotides downstream of the leader sequence with a bi-directional encapsidation process, Blumberg *et al* (1991) speculated that the process of genome encapsidation in paramyxoviruses could proceed from a nucleation site not necessarily located within the leader sequence. The role of the terminal sequences in replication has been demonstrated in influenza virus (Luytjes *et al*, 1989), HRSV (Collins *et al*, 1991) and Sendai virus (Park *et al*, 1991). The negative sense strand of chloramphenicol transferase (CAT) gene flanked by 5' and 3' NS gene regions was transcribed from a T7 RNA polymerase promoter *in vitro* and then added to purified influenza virus polymerase complex. The resulting RNP complex was transfected into tissue culture cells followed by infection of these cells by influenza virus. CAT gene expression was measured by the conversion of radiolabelled chloramphenicol. Changing just 3 amino acid residues within the 5' end of the RNA resulted in the loss of detectable CAT activity. A variation of this protocol in determining the importance of the 3' and 5' regions of Sendai virus (Park *et al*, 1991) and HRSV (Collins *et al* 1991) in replication has also been used. In these cases RNA of CAT gene flanked by 3' and 5' end

sequences transcribed from a T7 RNA polymerase promoter was directly used to transfect virus infected cells. For HRSV replacing the 3' leader sequence with the 5' trailer sequence that was 91% or 73% identical to the leader sequence in its first 11 or 22 nucleotides had no effect on CAT activity. However the addition of 11 heterologous nucleotides to the 3' leader abolished CAT gene expression indicating the conservation of a functionally important domain that is inactive when in an internal location. The removal of either 3 nucleotides from the 3' terminal region had no effect on CAT activity. For all three virus systems, the CAT reporter RNA was found to be packaged into virions by following passage of the virus present in the supernatant of these transfected-infected cells (Luytjes *et al*, 1989, Park *et al*, 1991, Collins *et al*, 1991). The presence of the complementary sequences at the 3' leader and 5' trailer regions may also be involved in choosing between transcription and replication. For example, if the viral polymerase binding site is located within the complementary regions, then the formation of an RNA duplex within this region might signal start of replication while a linear strand signals initiation of transcription or vice versa.

The viral polymerase carries out the functions of both transcription and replication of the genome. The switch between transcription and replication is thought to be achieved by an association between the polymerase, phosphoprotein and nucleocapsid protein. Replication of pneumoviruses occurs within the cytoplasm of the cell since HRSV has been shown to replicate in enucleated cells (Pringle, 1987). Replication is achieved by copying viral genome into full-length positive sense RNA, which is then used as a template for synthesis of the full-length negative, genomic, sense RNA.

1.6. PVM GENOME :

Harter and Choppin (1967) demonstrated that the genome of PVM is a single stranded RNA molecule. No effect on replication of PVM was observed in the

presence of 5-fluoro-2'-deoxyuridine (FUDR), 5-bromo-2'-deoxyuridine (BrdU) or 5-iodo-2'-deoxyuridine (IUDR), which are inhibitors of DNA synthesis, indicating the PVM genome to be an RNA molecule. Haematoxylin and eosin staining revealed eosinophilic cytoplasmic inclusions within PVM infected cells. Acridine orange which stains double stranded DNA or RNA green and single-stranded DNA or RNA flame red showed the genome of PVM to be single-stranded since red colouration was observed close to the eosinophilic cytoplasmic inclusions found in PVM infected cells. The negative polarity of the PVM genome is inferred from that of HRSV and the detection of specific mRNA species in infected cells (Chambers *et al*, 1990a). Genomic HRSV RNA was identified by the failure to amplify RNA in the presence of the protein synthesis inhibitor, cyclohexamide. Under these conditions viral replication is inhibited because of its reliance on continued viral protein synthesis whereas viral transcription proceeds as normal. The genomic RNA was used in specific annealing experiments which demonstrated that the genome of HRSV is of negative polarity (Huang and Wertz, 1982).

Chambers *et al* (1990a) have cloned and sequenced nine of the genes encoded by the approximately 16,000 nucleotide long single stranded, negative sense RNA genome of PVM. Table 1.1 shows the starts and ends of the known PVM genes which show similarity to those of HRSV (Chambers *et al*, 1990b). The conservation of the gene starts within each genome are presumably a requirement for recognition by the viral polymerase during transcription of each genetic unit. HRSV shows considerable conservation in its gene start sequences but not in its gene ends whereas PVM gene starts and ends are fairly well conserved (Chambers *et al*, 1990b). However no correlation between gene signal (start and ends) and level of transcription has been observed.

Gene	Gene start		Gene end	
	PVM	HRSV*	PVM	HRSV*
1C	AGGACAAGU	GGGGCAAAU	UAGUUAAUU	UAGUUAAUAU
1B	AGGAUAAAU	GGGGCAAAU	UAGUUAUAG	UAGUAAUUU
N	AGGAUAAAU	GGGGCAAAU	UAUUUAAUU	GAGUUAAU
P	AGGACAAAU	GGGGCAAAU	UAGUUAAUU	UAGUUAC
M	AGGAUAAAU	GGGGCAAAU	UAGUUAAAU	AAGUUAAU
1A (SH)	AGGAUAAGU	GGGGCAAAU	UAGUUAAC	UAGUUAAUU
G	AGGACAAAU	GGGGCAAAU	UAGUUAAUG	UAGUUACUU
F	AGGAUGAGU	GGGGCAAAU	UAGUUAAUU	UAGUUAUAU
M2 (22K)	**AGGACAAAU	GGGGCAAAU	UAGUUAUAU	UAGUUAUU
L	Unknown	***GGGACAAAA	Unknown	AUCUUAUAAU
Concensus	$AGGA^{C}/_{U}AA^{A}/_{G}U/_{G}$	$GGG^{G}/_{A}CAAA^{U}/_{A}$	$UA^{G}/_{U}UUA^{A}/_{U}$	$U/G/_{A}A/_{U}G/_{C}U^{U}/_{A}A$

Table 1.1 : Consensus sequence of PVM and HRSV gene starts and ends.

Nucleotide sequence of the known gene start and ends in PVM and HRSV and overall consensus sequence representing conservation in putative recognition sites for the viral RNA polymerase during transcription (adapted from Chambers *et al*, 1990b). * Gene start and ends for HRSV are shown for strain A2. **Two putative gene starts for the M2 (22K) protein of PVM are also found within the F-M2 intergenic region (Chambers *et al*, 1990b). *** The L gene start of HRSV is located within the M2 (22K) gene. Where two or more bases are indicated within the consensus sequences, the uppermost base represents that found predominantly.

The 'stop-start' model of transcription is thought to account for the decreased concentrations of mRNA species representing genes distal to the polymerase binding site located within the leader sequence at the 3' end of the genome. Upon reaching the polyadenylation signal of the nearest gene, the polymerase dissociates from the RNA molecule and either restarts transcription from within the 3' leader sequence or can re-initiate mRNA synthesis on the viral RNA molecule at the consensus sequence of the next gene. A consequence of this type of 'stop-start' transcription is decreasing amounts of mRNA synthesised for each gene, the further the gene is located from the 3' polymerase binding site. Chambers *et al* (1990a) have used this feature to determine the gene order of PVM from the abundance of cDNA clones of each gene, which was subsequently confirmed by sequencing the gene junctions (Chambers *et al*, 1990b).

Fig. 1.2 shows the gene order of PVM as determined by Chambers *et al* (1990a) in comparison to that of morbilliviruses and paramyxoviruses. The gene order of morbilliviruses and paramyxoviruses is arranged so that the genes encoding structural proteins involved in replication and transcription (nucleocapsid and phosphoproteins) are located adjacent to the polymerase binding site and thus would be transcribed first. The envelope glycoprotein genes, which encode the proteins responsible for fusion and attachment activities of the virus, are located 5' of the matrix (M) protein gene and precede transcription of the viral polymerase gene. Transcription of the viral polymerase gene immediately following infection is probably not necessary since the enzyme is associated with the infecting nucleocapsid and sufficient amounts would be present to initiate replication/transcription.

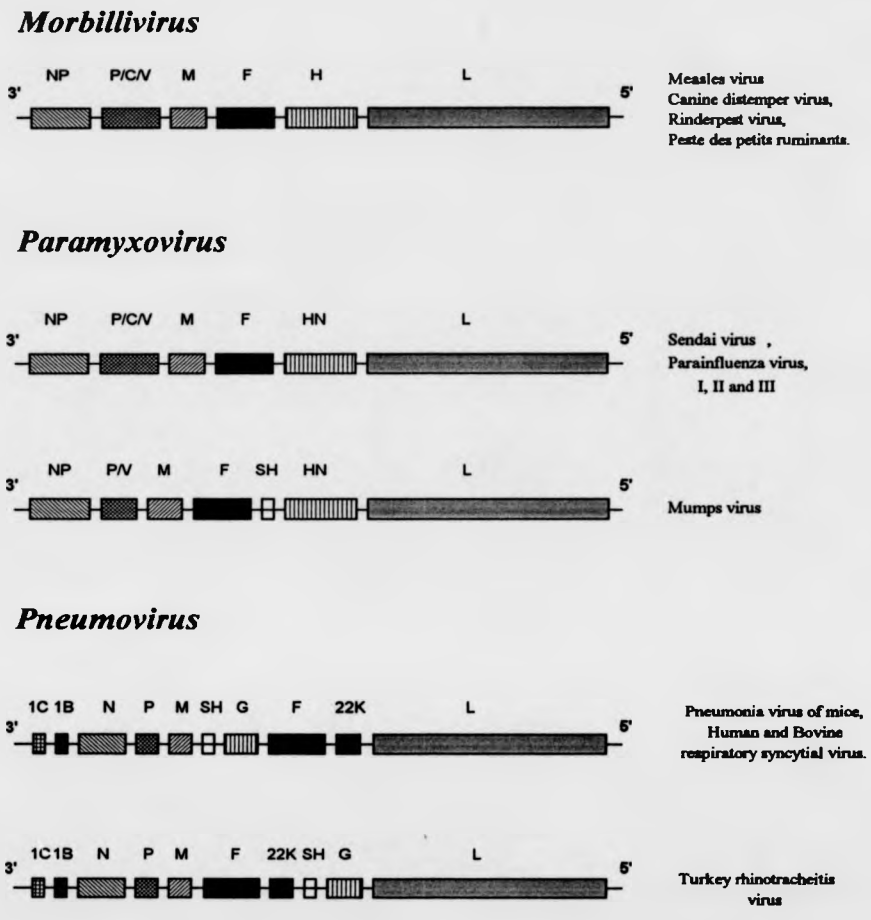


Fig. 1.2 : Gene order of *Paramyxoviruses*.

Gene order of PVM strain 15, as determined by Chambers *et al* (1990b), compared to that of morbilliviruses and paramyxoviruses. The gene order of TRTV is different to that of HRSV and PVM.

The genome of PVM, like HRSV, codes for genes unique to the pneumovirus genus. Unlike morbilliviruses and paramyxoviruses the structural genes encoding the nucleocapsid (N), phosphoprotein (P) and matrix (M) proteins are preceded by two small genes encoding non-structural proteins designated 1C and 1B. The M gene precedes the SH (1A) gene followed by the G (attachment) and F (fusion) genes. The M2 (22K) gene precedes the viral polymerase gene (L), but unlike HRSV where the start of the L gene lies within the M2 gene, no such gene start for PVM L gene is found within the M2 gene (Easton and Chambers, personal communication). The third pneumovirus, TRTV, is unusual in that the gene order is different to that of HRSV and PVM. The F and 22K genes are found to precede the SH and G genes (Ling *et al*, 1992, Yu *et al*, 1992), indicating a gene order similar to those of paramyxoviruses and morbilliviruses. Functions of PVM proteins can be assigned primarily by comparison with available data for similar proteins of other pneumovirus, morbillivirus and paramyxovirus members, and to some degree by analogy to proteins of members of the rhabdovirus genus, particularly VSV for which most information is available.

1.6.1. Non-structural proteins (1B and 1C) :

The genes encoding these proteins are unique to the pneumoviruses and are located closest to the 3' end of the genome RNA. Correspondingly, they are transcribed most efficiently of all the genes. The 1B and 1C genes of PVM encode polypeptides of 156 and 113 amino acids respectively (Chambers *et al*, 1991) which have a low or no homology with their HRSV counterparts. Although the polypeptides of these genes have been observed on polyacrylamide gel electrophoresis (Cash *et al*, 1979, Ling and Pringle, 1989a) they have not been not functionally characterised. The high abundance of their mRNA's and the absence of the equivalent polypeptides in virions of HRSV (Huang *et al*, 1985) suggest a role in viral replication rather than as

structural components of the virion. In pulse chase experiments, the 1B protein had a half-life of approximately 1 hour. However, turnover of the 1B protein was not observed when expressed from a baculovirus expression system (J. Evans, personal communication.). Cherrie *et al* (1992) have shown a high cytotoxic T cell (CTL) response to the 1B protein. Proteins expressed from recombinant vaccinia virus tested positive against serum from six of nine volunteers for CTL response against the 1B protein whereas only a low response to the 1C protein (one of nine) was observed. Under reducing PAGE conditions a polyclonal serum directed against the HRSV 1C protein immunoprecipitated a protein of approximately 13K whereas under non-reducing conditions, two proteins were observed at approximately 13K and 26K regions. These species may represent intra- and inter- disulphide bonded 1C aggregates respectively (J. Evans, personal communication.). Similar results have also been observed for the HRSV 1B protein in which multiple bands were observed under the non-reducing conditions. Additionally, neither the 1C or 1B proteins co-precipitated from 'pelleted' HRSV from infected supernatants, indicating these proteins to be non-structural (J. Evans, personal communication.). However, the HRSV 1C and M proteins were immunoprecipitated from HRSV infected cells with either an anti-M MAb or an anti-1C peptide serum indicating an association between the M and 1C proteins of HRSV (J. Evans, personal communication.).

1.6.2. Ribonucleocapsid proteins :

The function of the viral nucleocapsid proteins involved in transcription and replication of viral genome can be put into context following a brief description of the interaction of each of the components. It is generally thought that the N protein provides a template by binding to the viral RNA upon which the viral polymerase (L protein) subsequently acts. The P protein is thought to play a role in modification of the viral polymerase function from replication to transcription possibly in association with the N protein. The importance of the N, P and L proteins in viral replication is

exemplified by the work of Pattnaik and Wertz (1990) using VSV. Co-expression of the VSV N, P and L proteins *in vitro* was achieved by the use of recombinant vaccinia virus expressing the T7 RNA polymerase. Super-infection with non-replicative VSV defective interfering (DI) particles resulted in replication and amplification of the DI genome exceeding that observed in the presence of helper virus. Pattnaik and Wertz (1990) suggest that in the presence of helper virus both the helper virus and DI genomes compete for the newly synthesised NS, N and L proteins whereas in the *in vitro* system the requirement for such competition is reduced or eliminated resulting in increased replication/amplification of the DI genome. No DI replication was observed if either of the N, P or L proteins was omitted. Titration of each protein for optimal DI amplification provided evidence that an increase or decrease in concentrations of either the L or NS proteins but not the N protein resulted in decreased amplification of the DI genome. Following immunoprecipitation from cells transfected with the N protein alone radioactive heterodisperse RNA was observed. This RNA was found to be derived from the vaccinia virus infected host cells and bound to the expressed N protein. Co-expression of the N protein with either the NS or L protein or both proteins abrogated such non-specific binding. The results show that although the N protein is capable of binding RNA, specificity for VSV RNA is seen only in the presence of the NS and L proteins (Pattnaik and Wertz, 1990). Furthermore, co-expression of N, P, L, M and G proteins resulted not only in DI replication but also in assembly, packaging and release of DI particles into the cell culture medium (Pattnaik and Wertz, 1991). In cells co-expressing the NS, N and L proteins and either the M or G protein, subsequent infection with DI particles did not result in maturation of the DI particle and no progeny DI particles were detected in the cell culture medium. Thus both the M and G proteins appear to be crucial in the maturation of VSV and suggests an interaction between the M and G proteins during viral maturation (Pattnaik and Wertz, 1991).

1.6.2.1. L PROTEIN :

The L gene encodes the major component of the RNA dependent RNA polymerase responsible for viral transcription and replication. As yet the L gene of PVM has not been sequenced. Since it is located the furthest from the 3' polymerase binding site, any mRNA species representing this gene would be found in minimal amounts. However, Ling and Pringle (1989a) observed a polypeptide with a molecular weight of 191K on SDS-PAGE gels which was thought to be the L protein. Assuming this size is correct it is estimated the L protein of PVM to be approximately 1750 amino acids long. This figure is smaller than that of HRSV whose L protein is 2165 amino acids long (Stec *et al*, 1991), and is also smaller than the L proteins of other paramyxoviruses. However the resolution of SDS-PAGE gels used for analysis of PVM specific polypeptides would not have given an accurate estimation of the L protein size. It is possible for the L protein to have been up to 200K in molecular weight , which would be closer to that of HRSV. Unlike HRSV or BRSV where the start of the L gene is located within the M2 (22K) gene (Stec *et al*, 1991, Zamora and Samal, 1992), no open reading frame is found within the M2 gene of PVM or the G gene of TRTV (the gene order of TRTV is different to that of HRSV and PVM) which could represent the start of the L gene in these viruses (Andrew Easton - personal communication, Yu *et al*, 1992, Ling *et al*, 1992). The VSV viral polymerase has been shown to possess kinase activity required for further phosphorylation of a phosphorylated form of bacterially expressed VSV NS protein (Barik and Banerjee, 1992) and is discussed in more detail later.

1.6.2.2. N AND P PROTEINS:

The nucleocapsid protein (N) of PVM is 1215 nucleotides long encoding a protein of 393 amino acids. The predicted molecular weight of the N protein is 43,141 which is in close agreement with the 41K molecular weight observed

on SDS-PAGE by Ling and Pringle (1989a) and Cash *et al* (1979). The overall homology between the N proteins of PVM and HRSV is 60% although this homology varies from 38% for the first 150 amino acids to 74% for the remaining amino acids of these two viruses. Within some regions (amino acids 245 to 315) this homology increases to 97% (Barr *et al*, 1991). This high degree of homology probably accounts for the cross-reactivity between PVM and HRSV N proteins observed by Ling and Pringle (1989a) and Gimenez *et al* (1984). Although the N protein of PVM (and pneumoviruses in general) are smaller than those of morbilliviruses and paramyxoviruses Barr *et al* (1991) have identified structural features common to these proteins, which may be relevant for the function of the protein.

The phosphoprotein (P) of PVM is 903 nucleotides long encoding a polypeptide of 295 amino acids with a predicted molecular weight of 29K, which is smaller than the 41K polypeptide observed on SDS-PAGE gels by Ling and Pringle (1989a) and identified by Chambers *et al* (1990a) as the P protein in hybrid arrest translation experiments. In *in vitro* translation systems this protein migrates with an apparent molecular weight of 39K. The discrepancy between predicted and observed molecular weights can not be accounted for by phosphorylation of the P protein alone but could be a result of failure of these acidic proteins to interact with SDS and also the three dimensional conformation of the protein. Additionally, the PVM P gene can encode a second smaller polypeptide of 138 amino acids with a predicted molecular weight of 16K initiated from an overlapping open reading frame. This is somewhat reminiscent of the C proteins of several paramyxoviruses. The 39K P protein of PVM shares an overall homology of 28% with the P protein of HRSV and 35.6% with the P protein of bovine respiratory syncytial virus (BRSV). As was the case for the N protein the homology between the P proteins of PVM and HRSV rises to 64% in some regions of the protein. A low homology between the P proteins of pneumoviruses is not surprising since the NS protein of VSV, which is equivalent to the P proteins of

paramyxoviruses, demonstrated a low homology between different strains of the same virus. The NS protein of the Ogden strain of VSV was found to have a 30% homology with the same protein from New Jersey strain of VSV. However although the overall homology was low, the C termini of these proteins were well conserved with a homology of approximately 90% suggesting conservation of a functionally important domain (Gill *et al*, 1986, Banerjee, 1987).

By analogy to other paramyxoviruses it is generally considered that the viral polymerase interacts with the N and P proteins during transcription and replication. The precise mechanism by which the polymerase switches from its transcriptive to replicative mode of functions is as yet unknown although evidence suggests a possible role for the P protein. Caravokyri *et al* (1992) have attributed a mutation of a temperature sensitive mutant of HRSV to a single amino acid change in the P protein. At the non-permissive temperature no detectable viral protein was synthesised, probably due to the absence of mRNA species. These results suggest a role for the P protein in viral transcription. Evidence from studies of Sendai virus and NDV also indicate the crucial role of the P protein in viral transcription (Hamaguchi *et al*, 1983, Deshpande and Portner, 1985).

The Far-Western (West-Western) blot analysis technique has been used to study the interaction of the N and P proteins of PVM (J. Barr- personal communication). The technique involves immobilising PVM infected cell proteins to Immobilin™ polyvinylidene difluoride (PVDF) membrane using the Western blotting protocol after which the immobilised proteins are renatured in the appropriate buffer. Interaction of the proteins can then be assessed by incubation of the immobilised proteins with radiolabelled *in vitro* translated products followed by autoradiography. The extent of binding can then be assessed by densitometric analysis of the autoradiograph. Far-Western blot analysis of N- and C- terminally deleted N protein of PVM demonstrated an interaction with the P protein of PVM. The amino terminal half

of the N protein was shown to bind to immobilised P protein with a 4.3% affinity when compared to the binding affinity of full length N protein (100%). Similarly the C-terminal half of the N protein bound with 17.6% affinity to immobilised P protein. Removal of just twelve amino acids from the C-terminus of the N protein resulted in reduction of binding affinity to 68.4%. The results show that both the N and C termini are important and suggests that the overall structure of the protein is important (J. Barr- personal communication).

A number of independent domains located within the P protein were identified to be responsible for binding to the N protein (J. Barr- personal communication). Similar results for the binding between the NP and P proteins of Sendai virus have been observed. Removal of 30 amino acids from the C terminus of Sendai virus P protein abolished binding to immobilised NP protein (Homann *et al*, 1991) which support the findings of Ryan and Kingsbury (1988). Deletions within the Sendai virus NP protein considerably reduced the affinity of the NP protein for the P protein indicating that most of the NP protein was required for binding the P protein (Homann *et al*, 1991) . Deletions of the P protein of PVM at either the amino or carboxyl terminus totally abolished binding to the full length N protein, suggesting these regions were required for binding the N protein (J. Barr- personal communication). These results are also in agreement with the findings of Huang *et al*, (1984) and Lambert and Pons (1983) who observed an association between the P and viral nucleocapsids of HRSV. Garcia *et al* (1993) recently demonstrated the presence of N-P complexes in HRSV infected cells. Electron microscopic studies revealed the presence of cytoplasmic inclusions in which HRSV infected cells were stained with MAb's directed against the structural proteins. Inclusions were labelled with MAb's specific for the P, N and M2 (22K) proteins but not for the F, G or M proteins. Immunofluorescence staining with MAb's for N, P and M2 revealed the presence of cytoplasmic O-ring inclusions. Transfection of cells with N, P M2 genes, either alone

or in combination, followed by immunofluorescence and immunoprecipitation analysis showed interaction between the N-P and N-M2 (22K) proteins. In N and P co-transfected cells round aggregates within the cytoplasm were observed. Co-transfection of N and M2 or P and M2 showed diffuse cytoplasmic fluorescence whereas co-transfection with all three showed O-ring cytoplasmic inclusions similar to those found in HRSV infected cells, indicating an interaction between all three proteins. SDS-PAGE analysis of immunoprecipitations of transfected cells revealed presence of N-P and N-M2 but not P-M2 complexes. These complexes were only observed in infected or transfected cells but not when the proteins were co-synthesised in rabbit reticulocyte lysate (Garcia *et al*, 1993).

Fooks *et al* (1993) expressed measles virus nucleoprotein from a recombinant baculovirus in insect cells. Using electron microscopy they saw nucleocapsid like structures located within the nucleus and cytoplasm of these cells. The nucleocapsid structures were longer than those seen in measles virus infected cells and the buoyant density of these baculovirus expressed nucleocapsids was less than those observed from measles virus infected cells. This is consistent with the idea that the nucleocapsid structures in these insects cells lacked significant amounts of RNA. Similar nucleocapsid structures were observed in recombinant vaccinia virus infected cells expressing the measles virus N protein (Spehner *et al*, 1991). Thus, the N protein of measles virus is capable of interacting with itself and forming aggregates. Gombart *et al* (1993) used MAb N46 directed against measles virus N protein, which failed to react with P protein when expressed alone, to immunoprecipitate N-P complexes from cells co-expressing the N and P proteins in *in vitro* experiments. Again N-P complexes were only observed when both proteins were co-expressed. These results confirmed the findings of Huber *et al* (1991) who inferred an interaction between N and P proteins in measles virus following retention of the N protein within the cytoplasm, as observed by immunofluorescence, when co-expressed with the P protein, since the N

protein of measles virus can localise to the nucleus when expressed alone. Use of P protein deletions allowed them to attribute approximately 40% of the C terminal region of the P protein to binding to the N protein. Masters and Banerjee (1988a) demonstrated an association between the VSV nucleocapsid (N) and phosphoprotein (NS) in an *in vitro* expression system. They determined that N-NS complexes were only formed when both proteins were co-translated. Deletion of 21 amino acids of the C terminus of the VSV NS protein resulted in a decreased ability of the truncated protein to bind to RNP complexes (RNA-N) and had a minor effect on transcription (Gill *et al*, 1986). Horikami *et al* (1992) showed a requirement for NP-P and P-L complexes in Sendai virus. Replication of a Sendai virus defective interfering particle (DI) was supported in cells expressing either the NP and P or L and P proteins in the same cells. Furthermore they were able to immunoprecipitate NP-P and P-L complexes from these cells. A reduction or increase in either the NP or P proteins inhibited replication of the DI particle. Although no reason was attributed to the findings that an increase in the P protein expression reduced DI replication, it could be argued that an imbalance in the N:P ratio switches polymerase function from replication to transcription. Baker and Moyer (1988) demonstrated purified NP protein of Sendai virus was capable of supporting replication of purified DI nucleocapsids isolated from infected cells *in vitro*. However RNA synthesis was not observed when NP was added to detergent disrupted DI nucleocapsids. They concluded that the NP protein was interacting with another viral nucleocapsid protein presumably the P protein. Additionally, they interpreted their results to indicate that the NP protein is capable of RNA synthesis from pre-initiated templates only i.e. NP was involved in the elongation of already initiated RNA synthesis. Hamaguchi *et al* (1983) used detergent and salt protocols to strip NDV nucleocapsids of the P and L proteins. Following purification of the L and P proteins the effect of each upon nucleocapsid transcription was

investigated. In re-constitution experiments transcription activity was greater where both the L and P proteins than where the L and P proteins were used separately.

The interaction of the phosphoproteins of paramyxoviruses is complicated further by the synthesis of polypeptides from within the P gene either from 1) overlapping reading frames, 2) synthesis from non AUG initiation codons or 3) from insertion of non-templated nucleotides during transcription thus altering the reading frame as observed for the P/V genes of Sendai virus (Vidal *et al*, 1990), SV5 (Thomas *et al*, 1988), measles virus (Cattaneo *et al*, 1989) and human parainfluenza type 2 virus (Ohgimoto *et al*, 1990) and the P/D genes of HPIV3 (Galinski *et al*, 1992). The P/V genes of SV5 are unusual in that the V protein is coded from the first AUG whereas the P protein is obtained following insertion of non-templated bases. In all the other viruses mentioned above, it is the V protein that is accessed following insertion of non-templated bases. The latter is discussed more fully in the context of insertion of non-templated bases during transcription of the PVM G gene.

Sendai virus is unusual in that a number of polypeptides are synthesised from the P gene. The P protein is synthesised from the major ORF spanning most of the gene. The C polypeptide is 204 amino acids long and is synthesised from an AUG seven nucleotides downstream of the AUG encoding the 568 amino acid long P protein. Additionally, Sendai virus P gene encodes several other polypeptides -C', Y, Y', V, W and X (Curran and Kolakofsky, 1988). PIV1 (Matsuoka *et al*, 1991), canine distemper virus (CDV), NDV (McGinnes *et al*, 1988) and measles virus (Bellini *et al*, 1985, Alkhatib *et al*, 1988) also translate polypeptides from overlapping reading frames. The C' polypeptide is unusual in that it is translated from a non-AUG codon (ACG) located in-frame upstream of the initiation codon for the C polypeptide (Curran and Kolakofsky, 1988) in Sendai virus and from a GUG codon in PIV1 (Boeck *et al*, 1992). Both the C and C' polypeptides are non-structural. The Y and Y' polypeptides of Sendai virus are a result of internal initiation from within the C polypeptide coding

region whereas the V protein is translated from a separate mRNA species arising from insertion of a single non-templated G residue at a CCC (negative sense) insertion site located within the P gene. The V protein shares the first 316 amino acids in common with the N terminus of the P protein with a unique 68 amino acids cysteine-rich C terminus. The W polypeptide is a truncated form of the P protein which terminates after the addition of two extra codons following insertion of two, five or eight non-templated G residues at the same insertion site as the V polypeptide. Insertion of non-templated bases is thought to occur by the same method as polyadenylation of mRNA with stuttering at a run of C residues (Thomas *et al*, 1988, Vidal *et al*, 1990). Vidal *et al* (1990) have demonstrated this to be a feature of the viral polymerase during transcription since no non-templated insertions were found in vaccinia virus recombinants containing the full length P gene. The X protein of Sendai virus represents 95 amino acids of the C terminus of the P protein and is initiated from an internal AUG codon (Curran and Kolakofsky, 1988).

Absence of any mRNA synthesis in a ts mutant of HRSV due to lesion in the P protein at the non-permissive temperature, indicated a role for the P protein in providing specificity for the N-P complex to bind virus specific RNA (Caravokyri *et al*, 1992). This theory correlates with the hypothesis derived from the VSV results where *in vitro* translated N protein was shown to bind non-specifically to RNA (Masters and Banerjee, 1988a). However, co-translation of the N and NS proteins *in vitro* altered the specificity of the N protein to become specific for virus RNA (Masters and Banerjee, 1988b). These results complement those of Pattnaik and Wertz (1990) which have been discussed previously. Buchholz *et al* (1993) have demonstrated the ability of Sendai virus nucleocapsid protein (NP) to form nucleocapsid structures when expressed *in vitro* using the recombinant vaccinia virus T7 RNA polymerase expression system. Although these nucleocapsid structures were similar in appearance to Sendai virus nucleocapsid structure, they contained non-viral RNA. Deletion

analysis of the Sendai virus NP protein determined the N-terminal 399 amino acids to be involved in formation of nucleocapsid structures whereas the C-terminal (400 to 524 residue) region does not (Buchholz *et al*, 1993). Investigations of the non-structural polypeptides synthesised from the P gene in Sendai virus suggests a role for these polypeptides in transcription and replication. Curran *et al* (1991) looked at the effects of the L, NP and P/C polypeptides expressed *in vitro* on replication of a Sendai defective interfering particle (DIH4). All were found to be required for replication of the DI genome. Abolition of expression of the C polypeptide was found to have no effect on replication whereas co-expression of the V or W but not the X polypeptide reduced amplification of the DI genome. An increase in genome amplification was observed when the level of L protein was increased in the presence of the V protein (representing the N terminal half of the P protein) but not in the presence of a polypeptide representing the C terminus of the P protein. These results indicate that the N terminus of the P protein of Sendai virus is involved in binding RNA whereas the C terminus is involved in binding to nucleocapsid protein which support similar results for measles virus P protein (Huber *et al*, 1991) and VSV NS protein (Gill *et al*, 1986).

Curran *et al* (1992) have recently investigated the effect of Sendai virus NP, P/C and L proteins on mRNA synthesis. Extracts of transfected cells expressing, either separately or together, the NP, P/C and L proteins were reconstituted with purified RNP templates (NP:RNA). Extracts in which all proteins were co-expressed showed increased mRNA synthesis compared to those in which either one or two of the proteins was omitted. Extracts in which the P/C gene products were replaced with P gene only (C polypeptide not expressed) showed a five fold increase in mRNA synthesis. The C' polypeptide had an inhibitory effect on transcription. Reconstitution of extracts from co-expressed P and C polypeptides with L protein showed a strong inhibitory effect on mRNA synthesis whereas mixing of extracts from co-expressed C and L polypeptides with P protein showed only a minor inhibitory effect on mRNA

synthesis. These result indicate an interaction between the C and P protein affect transcriptional levels. Thus for Sendai virus it appears that the variation in the concentration of C and V proteins may be responsible for the switch from replication to transcription. Since the V protein is only found in all paramyxoviruses, except pneumoviruses, it could be a 'luxury' item.

A further complication is the finding that host cell proteins may be involved in transcription and replication of the virus. Barik and Banerjee (1991) purified a non-phosphorylated form of the VSV NS protein expressed in *E. coli*, designated P_0, which was found to be transcriptionally active when reconstituted with viral N-RNA template and L protein. However these studies were inconclusive since the preparations of L protein and N-RNA template were found to contain protein kinase activities, possibly due to kinases associated with the nucleocapsid. However, P protein synthesised using reticulocyte lysate was found to be phosphorylated (designated P_1) by kinases present in the reticulocyte lysate and was different to that expressed in *E. coli* (Barik and Banerjee, 1992). Extracts prepared from uninfected BHK cells were found to phosphorylate bacterially expressed P_0 to P_1. Both P_0 and P_1 were demonstrated to be transcriptionally inactive. Further phosphorylation of P_1 to a transcriptionally active P_2 form was achieved by the addition of purified VSV L protein and N-RNA template lacking any protein kinase activity, as measured by their inability to phosphorylate P_0 to P_1 (Barik and Banerjee, 1992). Further biochemical analysis determined that phosphorylation of the bacterially expressed NS protein to the P_1 form was achieved by host cell kinase, similar to caesin kinase. Modification of P_1 to P_2, the transcriptionally active form of the NS protein, was achieved by the addition of the L protein thereby demonstrating the role of the L protein in phosphorylation during maturation of the NS protein. Addition of the non-phosphorylated NS protein to just the L protein was not sufficient to restore transcriptional activity demonstrating the requirement for the host cell kinase and viral polymerase in a step-wise maturation

of the NS protein (Banerjee and Barik. 1992, Barik and Banerjee, 1992). Furthermore the P_2 form of the protein was found to undergo de-phosphorylation and re-phosphorylation in its phosphate groups attached by the L protein associated kinase activity. These results complement the findings of Chattopadhyay and Banerjee (1987) who demonstrated the importance of phosphorylation in transcription. Two conserved serine residues in the NS protein sequences of VSV were identified as important in viral transcription by the work of Gill *et al* (1986), where deletion of domain II (containing the serine residues) of the NS protein resulted in a marked reduction of transcription. Additionally a third non-conserved serine residue was located with this domain. Site directed mutagenesis of the these serine residues to alanine had considerable effect on transcription in reconstituted reactions. When either of the conserved serines were mutated, transcriptional levels were decreased by 75% of the activity of non-mutated NS protein whilst mutation of both the conserved serines reduced transcription by 90%. Mutation of the non-conserved serine only reduced RNA synthesis by 35%. When all three serines were changed, transcriptional levels fell by 92%. Thus the phosphorylation process of the NS protein appears to be involved in the regulation of viral transcription. The results presented above might suggest a role for phosphorylation to act as an 'on-off' switch for transcription although since complete inhibition of transcription following mutagenesis of the serine residues did not completely inhibit transcription (Chattopadhyay and Banerjee, 1987), phosphorylation of the NS protein might be thought of as a catalyst of viral transcription.

Another approach to the study of the role of the P proteins involves the use of monoclonal antibodies. Williams *et al* (1988) used three MAb's (2A2, 6D11 and 4F11) raised against purified VSV NS protein to investigate the role of the NS protein. Of the three MAb's used, two were inhibitory in *in vitro* transcription reactions using solubilised whole virus against which each of the MAb's was tested. Two of the three

MAb's (2A2 and 6D11) inhibited transcription by 90% whereas MAb 4F11 inhibited transcription by 65%. Further investigation into the mode of action of MAb's 2A2 and 6D11 showed MAb 6D11 to inhibit leader RNA synthesis whereas MAb 2A2 did not, as observed by the presence of newly synthesised leader RNA during transcription of a VSV DI particle. Furthermore, MAb's 6D11 and 4F11 immunoprecipitated the NS protein from purified nucleocapsid preparations whereas MAb 2A2 cross-reacted with nucleocapsid associated N protein. The mode of action of each MAb is determined from the fact that MAb 6D11 immunoprecipitated the NS protein with only background levels of the N protein indicating a dissociation of the NS protein following binding of the MAb whereas MAb 4F11 immunoprecipitated all the components of the nucleocapsid, indicating a binding to the NS protein in which the nucleocapsid remains intact following MAb binding. MAb 2A2 was shown to cross-react with a greater affinity for the N protein rather than the NS protein implying inhibition of transcription was achieved by binding to both the N and NS proteins. Although the use of MAb's may provide some insight into the mode of action of the antibodies themselves, they do not provide directly provide evidence for the exact role of the nucleocapsid associated proteins.

The importance of the V and C proteins in replication and transcription is unknown. Polypeptides synthesised from internal initiation of the PVM P gene have been observed both *in vitro* and *in vivo* although they have not been characterised as yet (J. Barr, personal communication). However, it has been suggested that the non-structural proteins of pneumoviruses may play a similar role (Collins, 1991). Although no conclusive evidence exists for the factor responsible for the switch from replication to transcription function of the viral protein, evidence to date suggests the P protein plays a crucial role in viral transcription and replication. Combining results available for the different viruses, and discussed above, it could be argued that although the N protein is capable of associating with the viral genome, the interaction of the N and P

proteins provide specificity for the N protein to bind viral specific RNA and possible interactions between the P and L proteins may switch the function of the viral polymerase from replication to transcription.

1.6.3. Membrane associated non-glycosylated proteins :

1.6.3.1. MATRIX (M) PROTEIN :

The matrix protein of PVM is 257 amino acids long and has some homology with its counterparts in HRSV and TRTV (Andrew Easton - personal communication). The predicted molecular weight of 28,312 is in close agreement with the 27.3K polypeptide identified to be the M protein following immunoprecipitation by Ling and Pringle (1989a) and hybrid arrest translation *in vitro* (Chambers *et al*, 1990a). The predicted protein shares a hydrophobic nature with its counterparts in other paramyxoviruses. The paramyxovirus matrix protein is thought to play a crucial role in virus assembly through interaction between the cell membrane and the viral nucleocapsid (Peeples, 1991). This view is supported by several lines of evidence.

Membrane association of the M protein following reconstitution with liposomes in Sendai virus (Caldwell *et al*, 1986) and NDV (Faaberg and Peeples, 1988) indicates a direct interaction with membrane components. NDV M protein was purified from virions by monoclonal antibody affinity chromatography and prevented from forming aggregates by the addition of bovine serum albumin (BSA). *In vitro* synthesis of liposomes was achieved by the reaction of cholesterol, L-α-phosphotidylcholine and a third lipid possessing either a positive, neutral or negative charge. In all cases the M protein was found to interact with the liposomes as analysed by reactivity of the M-liposomes complex with three M protein specific MAb's suggesting a non-electrostatic interaction between the two components. Furthermore the ability of M protein aggregates which lacked BSA to bind to the liposomes eliminated a role for BSA in mediating the interaction. Since the M proteins of

paramyxoviruses are hydrophobic, a non-electrostatic interaction with liposomes could be mediated by hydrophobic forces (Faaberg and Peeples, 1988). The M protein specific MAb's failed to prevent binding to the liposomes indicating a role for hydrophobic forces since the hydrophobic domains of the M protein are unlikely to be immunogenic. The possibility exists that these hydrophobic domains may be involved in interactions with the nucleocapsid or with membrane bound glycoproteins (Faaberg and Peeples, 1988). Chong and Rose (1993) demonstrated an interaction between purified VSV M protein and cell membranes and between the membrane bound M protein and the RNP core. No interaction of membrane bound G protein and the RNP core was observed. Interactions between the M and NP proteins of Sendai virus were observed by Homann *et al* (1991) in experiments using immobilised proteins to investigate protein-protein interactions. However, they also noted interactions between the M protein and non-viral (host) proteins. Ogden *et al* (1986) have demonstrated interaction between the VSV M protein and unilamellar vesicles and between the M protein and VSV RNP. Unlike Faaberg and Peeples (1988), Ogden *et al* (1986) could only show binding between the VSV M protein and negatively charged unilamellar vesicles. VSV RNP's were found to complex with these vesicles with a greater affinity in the presence of the M protein. Interestingly, complexed RNP-M bound with greater affinity to vesicles not already complexed with the M protein. One explanation suggested is the possible reduction in binding sites on M protein saturated RNP's. These results suggest that binding of M protein to membrane does not require the presence of viral glycoproteins within the cell membrane.

In temperature sensitive mutants of NDV in which the ts lesion lies in the M protein, reduction in plaque formation and haemolysis is attributed to the absence or reduction in incorporation of the F protein at the non-permissive temperature (Peeples and Bratt, 1984). Normal incorporation of F protein within membranes was found in revertants of these ts mutants in which the M protein had

reverted to normal SDS-PAGE migration. This suggests an interaction between the M and F proteins of NDV.

Analysis of temperature sensitive mutants arising from mutations found in the M protein by Caravokyri and Pringle (1991) using nucleotide sequencing and SDS-PAGE analysis identified an abnormal M protein in a temperature sensitive mutant of HRSV causing an increased degradation of the M protein at the non-permissive temperature, thus inhibiting viral assembly.

Interestingly, the M protein of measles virus could account for a change in the course or type of disease caused by this virus. Analysis of the M proteins of the Nagahata strain, an acute measles virus progenitor, and the Biken strain, isolated from a patient suffering from SSPE, showed differing abilities to bind to viral nucleocapsids. Whereas the M protein from the Nagahata strain was able to bind to the nucleocapsid as expected, that of the Biken strain was retained within the cytosol and did not bind to the viral nucleocapsid (Hirano *et al*, 1992). Earlier, nucleotide sequencing and polypeptide analysis by SDS-PAGE of a measles virus isolate from a patient afflicted with SSPE confirmed the absence of M protein synthesis resulting in the non-evagination of the virus. In this case an amber stop codon was found 36 nucleotides (12 triplets) into the coding region of the M protein (Cattaneo *et al*, 1986).

Additionally, by analogy to VSV (Ogden *et al*, 1986, Black and Lyles, 1992, Black *et al*, 1993) and Influenza A virus (Ye *et al*, 1989), the matrix protein may inhibit host cell transcription in a process which requires localisation of viral M protein to the cell nucleus. Inhibition of transcription was monitored by turnover of radiolabelled chloramphenicol incubated with cell extracts following co-transfection with recombinant SV40 based expression vectors containing the VSV M protein and CAT reporter gene (Black and Lyles, 1992). Northern blot analysis confirmed inhibition to be at the transcriptional level following a reduction in the observed CAT

mRNA concentrations in transfected cells (Black and Lyles, 1992). Digestion with trypsin resulted in the removal of 43 amino acids from the N terminus of the VSV M protein. This truncated protein was shown to have a reduced inhibitory effect on transcription *in vitro* (Ogden *et al*, 1986). Li *et al* (1993) also demonstrated that M protein expressed and purified from insect cells was capable of inhibiting transcription when reconstituted with VSV nucleocapsids. Li *et al* (1993) also expressed the VSV M, N and G proteins in insect cells using a baculovirus expression system. Electron microscopy demonstrated that the M protein alone was capable of initiating budding at the cell membrane. These results support those of Blondel *et al* (1990) who used a ts mutant (ts lesion attributed to a mutant M protein) to demonstrate that the M protein is responsible for cytopathic effect (CPE) in VSV infected cells. At the non-permissive temperature, cells transfected with ts M protein appeared normal whereas those at the permissive temperature rounded up and became detached. The results of Li *et al* (1993) indicate that no interaction between the glycoproteins and M protein of VSV is required to cause evagination, although it is not known whether host cell proteins present at the cell surface of insects cells have similar characteristics to the virus glycoproteins. These findings are in contrast to those of Pattnaik and Wertz (1991) who showed maturation of a non-replicative DI particle did not occur in an *in vitro* replication system in which transfected cells co-expressing the N, NS, L and either the M or G protein of VSV prior to super-infection with DI particles. These results suggest an interaction between the M and G proteins during VSV maturation. Since the M protein of pneumoviruses is considerably smaller than those of morbilliviruses and paramyxoviruses, these functions may have been split between the M and M2 proteins. Recently, Black *et al* (1993) have identified the VSV M protein domains involved in transcription inhibition and membrane attachment. Two mutant M proteins were engineered using standard molecular biology techniques. The first mutation involved changing of a Met residue to Arg at position 51 of the M protein of VSV

strain Orsay. This mutation, designated 082M, altered the M protein to that found in the ts mutant ts082 which was studied by Coulon *et al* (1990). The M protein of this mutant had been shown to produce a reduced CPE. The second mutation, designated MN1, was the deletion of the amino acids 4 to 21 in the M protein of VSV strain San Juan. This deletion was similar to that of Blondel *et al* (1990) who demonstrated that deletion of the lysine rich N terminus did not affect the ability of the truncated M protein to induce CPE. Inhibition of transcription was monitored by the method of Black and Lyles (1992). Cells co-expressing the CAT reporter gene and 082M showed minimal inhibition of transcription whereas those expressing the CAT reporter gene and MN1 showed marked reduction in transcription as measured by the ability of transfected cell extracts to turnover radiolabelled chloramphenicol. Ability of each mutant M protein to associate with cell membranes was investigated by the ability of each mutant to complement growth at the non-permissive temperature of a ts mutant (ts023) whose ts lesion has been attributed to the M protein. Here, 082M was found to complement growth of the ts mutant, as measured by plaque assays whereas MN1 did not complement growth. These results demonstrate the sites responsible for transcription inhibition and virus assembly to be separate and distinct. Additionally, each of the mutant proteins were found to localise to the nucleus as efficiently as the wild type M protein in immunofluorescence studies. Thus deletion of the lysine rich N terminus (MN1) is not involved in the nuclear localisation of the M protein (Black *et al*, 1993). Localisation of the NDV M protein to the nucleus have been observed in immunostaining studies (Peeples *et al*, 1992) and are similar to the findings of Lyles *et al* (1988), who used immunofluorescence to show localisation of the VSV M protein to the cell nucleus.

1.6.3.2. M2 (22K) PROTEIN :

The M2 (22K) protein is only found within members of the pneumovirus genus. The M2 protein of PVM encodes a polypeptide of 176 amino

acids with a predicted molecular weight of 20,165. The basic nature of the predicted polypeptide puts it in close agreement with the 23K or 24K polypeptides identified by Ling and Pringle (1989a) by two dimensional electrophoresis. The hydropathy profile resembles that of the M2 protein of HRSV and TRTV highly conserved between all three (Andrew Easton - personal communication). As yet the function of the M2 protein of pneumoviruses has not been established. The location of the protein within the virion has not been determined precisely although fluorescent antibody studies have shown it to be associated with internal structures of HRSV such as the N and P proteins (Routledge *et al*, 1987). Similar results were obtained by Garcia *et al* (1993) who demonstrated an interaction between the N and M2 protein by immunofluorescence and immunoprecipitation studies (discussed in the N and P protein section). This location is distinct from that occupied by the matrix (M) protein. These results agree with the findings of Samal *et al* (1993) who demonstrated an interaction between the M2 protein and nucleocapsid (N) protein of BRSV, following synthesis of the M2 protein in insects cells. Huang *et al* (1985) used detergent (triton) and salt dissociation methods to show an interaction of the M2 protein with the viral membrane. At high salt concentrations the M2 protein was solubilised whereas those associated tightly with the nucleocapsid (N, P and L proteins) were not, indicating the M2 protein to be membrane associated. However, whereas all of the M2 protein could be extracted from virion membranes, a small amount of the M protein remained membrane bound (Huang *et al* 1985). This could suggest, taking in to account the ability of the M2 protein to bind to the N protein, a possible involvement in the bridging of N and M proteins. Detergent and salt dissociation methods could disrupt a loose association with the nucleocapsid but not a tight association with the M protein. Thus the M2 protein would be solubilised and appear to be membrane bound. Additionally, Connors *et al* (1992) demonstrated a vaccinia virus recombinant expressing the M2 protein to be targeted by $CD8^+$ T cells since depletion of $CD8^+$ T

cells abrogated protection of immunised mice following infection with HRSV which had been observed previously (Nicholas *et al*, 1991). Nicholas *et al* (1991) observed a reduction in HRSV replication infected mice previously immunised with a vaccinia virus recombinant expressing the M2 protein.

Although, paramyxoviruses and morbilliviruses do not have an equivalent gene, it is possible that the role of their matrix (M) protein which is involved in the structure of the virion and possibly in regulation of transcription (Satake and Venkatesan, 1984, Ogden *et al*, 1986, Black and Lyles, 1992, Black *et al*, 1993) has been split in the pneumoviruses. A simpler model for the functioning of the M2 protein could be where the M2 protein acts as a 'glue' which binds the viral nucleocapsid to the M protein which in turn interacts with the envelope associated viral glycoproteins. This model presumes that the M2 and M proteins are not involved in regulation of transcription as is the case for VSV M protein.

1.6.4. Lipid envelope glycoproteins :

The major antigenic determinants of paramyxoviruses are the membrane-bound glycoproteins responsible for entry of viral nucleocapsid into the host cell. Attachment of the virus particle to the cell membrane is mediated by the attachment glycoprotein and fusion of viral and cell membranes, allowing entry of the nucleocapsid into the cell, is ascribed to the fusion (F) glycoprotein of these viruses. Since the fusion and attachment activities of VSV are integrated in the single surface glycoprotein, G, comparison with only the *Paramyxovirus* glycoproteins may be valid.

1.6.4.1. SH (1A) PROTEIN :

The SH gene of PVM is 390 nucleotides long, encodes a polypeptide of 92 amino acids with a predicted molecular weight of 9,582 for the unglycosylated protein and contains a large central hydrophobic region characteristic of pneumovirus SH proteins (A. Easton, personal communication). Ling and Pringle (1989b) identified

a 12K protein immunoprecipitations of PVM infected cells. The size of this polypeptide discounted it as being the F_2 subunit of the functionally active F protein. However, the size and membrane location of this 12K polypeptide could suggest it to be the SH protein of PVM. The SH protein is unusual in that an equivalent is found in all members of the pneumovirus genus and only in some members of morbilliviruses and paramyxoviruses such as SV5 (Hiebert *et al*, 1988) and mumps virus (Elango *et al*, 1989, Takeuchi *et al*, 1991). Although the function of the SH protein is not known, there is evidence supporting the membrane-bound association of this protein (Olmsted and Collins, 1989, Collins and Mottet, 1993). The SH protein of PVM contains 4 possible N-linked glycosylation sites. Olmsted and Collins (1989) demonstrated the existence of glycosylated and non-glycosylated forms of the SH protein of HRSV. Only a small fraction of the protein was glycosylated by the addition of N-linked glycosylation at either one or both of the two putative N-linked glycosylation sites located within the N- and C- termini of the protein. Similarly the 174 amino acid SH protein of TRTV possesses a N-linked glycosylation site within the C terminus of the protein indicative of an extracellularly located domain. This N-linked glycosylation site appears to be utilised since a mobility shift of *in vitro* transcribed and translated TRTV SH protein was observed in the presence of microsomal membranes (Ling *et al*, 1992). Indirect immunofluorescence of HRSV infected non-permeabilised cells demonstrated the membrane location of the SH protein and implied the C-terminus was exposed extracellularly since the anti-SH serum was directed against the C-terminus of the protein (Olmsted and Collins, 1989). Following trypsinisation of membrane bound SH protein, an SH C-terminal specific MAb was unable to recognise the protected fragment in immunoprecipitation analysis whereas a SH N-terminal specific MAb recognised this fragment. These results imply an extracellular location for the SH protein C-terminus and a cytoplasmic or membrane embedded location for the SH protein N terminus. Sequence analysis of the SH gene of HRSV strain 18537, the

prototype strain of HRSV subgroup B revealed 76% amino acid homology with that of strain A2, the prototype strain of HRSV subgroup A. The amino acid changes were located within the C-terminal part of the protein suggesting the SH protein may be subject to selective pressure from host immunity (Collins *et al*, 1990). The SH protein of SV5 has also shown to be membrane bound with the N-terminus of the protein located cytoplasmically (Hiebert *et al*, 1988). However, only 5 amino acids of the SH protein are exposed extracellularly in SV5. Thus the orientation of the SH proteins appear to be analogous to the attachment proteins of these viruses and may play a similar role. Collins and Mottet (1993) suggest the SH protein to be involved in either entry of the viral nucleocapsid into the cell or in maturation of the virions.

1.6.4.2. FUSION (F) GLYCOPROTEIN :

The fusion protein of paramyxoviruses is responsible for the fusion of viral and cell membranes during infection and also mediates fusion of the membrane of infected cells and adjacent cells allowing spread of the virus. The route chosen for nucleocapsid entry into the cell involves the fusion of viral and cell membranes at neutral pH in paramyxoviruses.

The F gene of PVM is 1657 nucleotides long encoding a polypeptide of 537 amino acids with a predicted molecular weight of 59K (Chambers *et al*, 1992). This is somewhat larger than the 49K molecular weight observed in hybrid-arrest experiments (Chambers *et al*, 1990a). The gene possesses nine non-coding nucleotides at the 5' end of the F gene with synthesis of the F protein starting at the first AUG and has a 3' non-coding region of 37 nucleotides (Chambers *et al*, 1992). The protein shares a 40% amino acid homology overall with its counterparts in HRSV and TRTV but only a 10% overall homology with the F proteins of NDV, SV5 or Sendai virus. The F proteins of HRSV have a high degree of conservation at the amino acid level not only between members of the same strain but also between members of subgroups

A and B, implying a low level of tolerance to changes within the F protein of these viruses (Collins *et al*, 1984, Elango *et al*, 1985, Baybutt and Pringle, 1987, Johnson and Collins, 1988).

The PVM and HRSV F proteins, like all other *Paramyxovirus* F proteins is synthesised as full length inactive precursor (F_0) which is cleaved during maturation by host cell proteases, into a functional protein consisting of disulphide-bond linked F_1 and F_2 subunits. The putative cleavage site for PVM F protein lies between amino acids 101 (Arg) and 102 (Phe). The F_2 subunit is 101 amino acids long whilst the F_1 fragment is 436 residues long. As is the case with the F_0 polypeptides of other paramyxoviruses, the F_0 polypeptide of PVM possesses a short N-terminal hydrophobic region that may act as the signal peptide which is subsequently cleaved during maturation, a C-terminal hydrophobic region thought to act as the transmembrane region and a fusion related hydrophobic region located at the N terminus of the F_1 subunit (Chambers *et al*, 1992). However F proteins of pneumoviruses are unusual in that their F_2 subunits have two cysteine residues whereas those of morbilliviruses and paramyxoviruses only have one. It is possible that both of these cysteines could be involved in disulphide bond formation with the F1 subunit. The significance of this upon the fusion function of the F protein is unknown. Ling and Pringle (1989b) were able to detect the F_0 and F_1 polypeptides of PVM on reducing SDS-PAGE gels but failed to detect the F_2 subunit possibly due to its small size and lack of glycosylation.

Although little homology is found between the F proteins of paramyxoviruses, Chambers *et al* (1992) have identified several structural features common to these proteins (Fig. 1.3). All the F proteins of paramyxoviruses are type 1 glycoproteins. The N-terminus of the F_2 fragment is predicted to consist of β-sheet or turn structures (F_2b) which are separated from C-terminus helical region (F_2h) by a region containing the conserved cysteine residue. The N-terminus of the F_1 fragment is

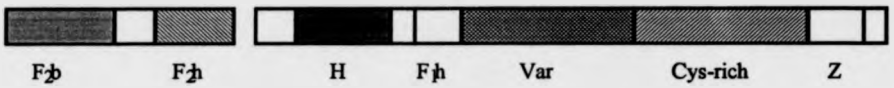


Fig. 1.3 : Conserved structural regions of *Paramyxovirus* fusion proteins.

Paramyxovirus fusion protein conserved structures identified and taken from Chambers *et al* (1992). Regions F_2b and F_2h correspond to the β-sheet and amphipathic α-helical regions of the F_2 subunit respectively. The F_1h region corresponds to the amphipathic α-helical region whereas the Var and Cys-rich regions correspond to the β-sheet regions of the F_1 subunit. H and Z regions correspond to the heptad repeat regions.

hydrophobic (fusion region) and thought to be involved in the fusion activity of the protein by localised disruption of the cell membrane during infection. The findings that synthetic peptides representing this region in Sendai virus inhibits cell fusion and virus penetration indicate this region is capable of interacting with lipid membranes (Richardson *et al*, 1986, Richardson and Choppin, 1983). A stretch of heptad repeat sequence, designated H, with a high α-helix structure is located next to this fusion region. The heptad repeat sequence consists of amino acids that have small side chains and are either neutral or hydrophobic occupying positions a and d in an amino acid sequence periodicity of a, b, c, d, e, f and g (Chambers *et al*, 1990c).

The H is followed by another hydrophobic region with a relatively low prediction for α-helix (F_1h) adjacent to a variable, cysteine-rich region. The cysteine-rich region contains the conserved cysteine residues with a number of proline residues indicative of β-sheet structures for this region (Chambers *et al*, 1992). Located C-terminal of the cysteine-rich region is another hydrophobic heptad repeat sequence

region, designated Z. This is analogous to the leucine zipper region of *Paramyxovirus* proteins identified by Buckland and Wild (1989). Finally, the extreme C-terminus of the F proteins contains a hydrophobic membrane spanning region followed by a small sequence that would be located cytoplasmically. Wathen *et al* (1989a) expressed the HRSV F protein lacking this C-terminus membrane anchor region from a baculovirus expression system and observed secretion of the protein into the cell medium. Also, expression of full-length F protein in insect cells resulted in limited cleavage of the protein implying lack of proteases in insects cells required for the correct processing of the proteins.

It has been demonstrated that changes at or around the cleavage site of the F_0 polypeptides influence not only cell or organ tropism but also the pathogenicity of the virus. The molecular basis for these observations are discussed in detail in chapter 3. Horvath and Lamb (1992) created F protein mutants with the H region, identified by Chambers *et al* (1992), of SV5. They determined that only those proteins transported to the cell surface showed fusion activity, as measured by the formation of syncytia. However, mutant F proteins in which the hydrophobicity of the region was increased did not produce increased fusion activity. Wang *et al* (1992) utilised the ability of transfected wild type NDV F protein to complement mutant NDV viruses in plaque assays. Sequence analysis of these mutants revealed changes located within the amphipathic α-helix, region H, identified by Chambers *et al* (1990c). The changes were found to be in positions b, c or g of the heptad repeat sequence identified by Chambers *et al* (1990c). Furthermore these changes were found to disrupt the α-helical structure of the H region resulting in defective transport of the F_1-F_2 protein to the cell surface. Anderson *et al* (1992) used recombinant vaccinia virus to express wild type full-length F protein and a mutant F protein of HRSV. The mutant F_0 protein was not cleaved into the F_1 and F_2 fragments. Nucleotide sequence analysis of the mutant F protein revealed four amino acids changes compared to the wild type full-length F

protein. Two of the changes were located within the F_2 fragment whereas the remaining two were located within the variable region of the F_1 fragment. Recombinant vaccinia virus expressing chimeric F proteins in which the mutant and wild type F_2 and F_1 regions were interchanged determined changes within the F_1 fragment to be responsible for abolition of F_0 cleavage. Expression of chimeric F proteins in which one or the other amino acid identified to be responsible for abolition of F_0 cleavage in the mutant F protein, showed a decrease, but not abolition, in the susceptibility to cleavage of the F_0 polypeptide (Anderson *et al*, 1992). These results demonstrate the influence of amino acid changes located away from the cleavage site and suggest a conformational change in which the cleavage site is inaccessible to host cell proteases during maturation when both amino acid changes are present. However, reversion of just one of these amino acid changes results in limited access to the cleavage site.

From these analyses it is clear that several distinct regions of the *Paramyxovirus* fusion protein are essential for the full functioning of the protein and that alteration to one or more of these regions can have profound effects on activity. This may be the reason for the apparent lack of sequence variation in fusion proteins of different HRSV strains and the conserved structural features described above.

1.6.4.3. ATTACHMENT (G) GLYCOPROTEIN :

The characteristics of the attachment glycoproteins was responsible for the initial classification of *Paramyxoviruses* into three genera. It is believed the haemagglutinin and neuraminidase activities of these viruses may have resulted from the linking of the HA and NA genes of an influenza-like ancestral virus (Blumberg *et al*, 1985, Morrison and Portner, 1991). Although there is little homology between the *Paramyxovirus* attachment proteins, they all belong to the class 2 group of glycoproteins. Characteristically, these proteins have a hydrophobic region at the N-

terminus which functions as the transmembrane region. No large hydrophobic regions are found at the C terminus that would be capable of acting as the membrane anchor. This N-terminal hydrophobic region is also thought to contain the signal sequence that is recognised by the signal recognition particle (SRP). Thus, in contrast to the fusion glycoproteins, the C terminus of the attachment glycoproteins is located extracellularly with the N terminus located intracellularly. Antigenic studies with anti-G protein monoclonal antibodies have shown the existence of two major subgroups (A and B) of HRSV based on differences in antigenicity (Mufson *et al*, 1985; Anderson *et al*, 1985).

The G protein of HRSV is unusual in that it is considerably smaller than its counterparts in paramyxoviruses and morbilliviruses. Whilst those of Sendai virus, SV5, NDV, measles virus, mumps virus and PIV3 are 576 (Blumberg *et al*, 1985), 565 (Hiebert *et al*, 1985), 577 (Millar *et al*, 1986, Jorgensen *et al*, 1987), 617 (Alkhatib and Breidis, 1986), 582 (Waxham *et al*, 1988) and 572 (Elango *et al*, 1986a) amino acids long, respectively, those of HRSV strains are either 298 or 292 amino acids long depending upon the subgroup (Wertz *et al*, 1985, Satake *et al*, 1985, Johnson *et al*, 1987b, Sullender *et al*, 1990, 1991, Sullender and Wertz, 1991). The G gene of PVM strain 15 is 1330 nucleotides long and encodes a polypeptide of 363 amino acids with a predicted molecular weight of 39,836. The major polypeptide, like that for HRSV G proteins, is transcribed from the second and not the first initiation codon. The haemagglutination activity of PVM resides on the G protein since a MAb that inhibited haemagglutination bound to the PVM G protein (Ling and Pringle, 1989b). The larger size of the PVM G protein when compared to that of HRSV could be a result of the acquisition of the region responsible for haemagglutination. The G protein of PVM is of similar length to the 391 residue attachment protein of TRTV. However, unlike PVM and HRSV this protein is transcribed from the first available initiation codon. Since the gene order of TRTV is different to that of HRSV and PVM where the G gene is located further downstream than in HRSV and PVM it is possible that the first

initiation codons in the G genes of HRSV and PVM are involved in down-regulation of G gene expression since most of the ribosomes would initiate translation from the first available AUG (Kozak 1986). In TRTV the requirement for such a down-regulating AUG is probably not required since the position of the G gene in TRTV is located further downstream than in PVM and HRSV and since genes are transcribed in a decreasing amount the further downstream they are located, expression of the G gene would already have been down-regulated. Since the pneumovirus attachment proteins are considerably smaller than those of paramyxoviruses and morbilliviruses and lack any sequence or structural homology it would be prudent to infer data for the PVM G protein from that available for the HRSV G protein.

Levine *et al* (1987) used two rabbit anti-G monospecific antisera to demonstrate the G protein of HRSV to be the attachment protein. One of the antisera was raised against purified G protein and the other against recombinant vaccinia virus expressing the HRSV G protein. Both antisera inhibited attachment of the virus to HeLa cells *in vitro*. Accordingly, it is thought the G protein of PVM serves a similar function. Ling and Pringle (1989b) demonstrated that the G protein of PVM was both N- and O-link glycosylated. In immunoprecipitation studies four G-related polypeptides in PVM infected cells were identified. Two were glycosylated forms with molecular weights of 76.4K and 62.0K. Appearance of the 76.4K glycosylated form preceded the appearance of the 62.0K form by 300 mins in pulse chase experiments. The appearance of the 62.0K form coincided with a reduction in intensity of the 76.4K form, suggesting that the 62.0K glycoprotein was derived form the 76.4K protein. Polypeptide synthesis in the presence of tunicamycin or digestion of the 76.4K and 62.0K proteins with glycopeptidase F, which removes N-linked carbohydrates, resulted in the reduction of observed molecular weight to 58.4K and 48.2K respectively, thus demonstrating the presence of N-linked sugars. The two forms of proteins synthesised in the presence of tunicamycin appeared to be processed normally, indicating that N-

linked sugars were not required for the correct processing of the G protein. However, digestion of either the 76.4K, 62.0K, 58.4K or 48.2K proteins with a mixture of endo-α-N-acetlygalactosaminidase and other exoglycosidases considerably reduced the observed molecular weights, demonstrating extensive O-linked glycosylation. The unglycosylated form of the 76.4K and 62.0K proteins were determined to be 39.6K and 37.2-44.8K respectively. Approximately 60% of the apparent 84,000-90,000 molecular weight of the mature HRSV G protein is contributed by carbohydrate, the majority of which is O-linked since the G protein is glycosylated in the presence of tunicamycin, an inhibitor of N-linked glycosylation. This is consistent with the high proportion of serine and threonine residues, the sites for O-linked glycosylation, representing 30% of the total amino acid content of the PVM and HRSV virus G proteins (Wertz *et al*, 1985, 1989, Fernie *et al*, 1985, this thesis). The G protein of TRTV also has a high serine, proline and threonine content of 10%, 14% and 7% respectively, though the level of glycosylation has not been established (Ling *et al*, 1992).

Collins (1990) demonstrated that the extracellular domain of HRSV G protein contains all the information necessary for correct O-glycosylation processing. Chimaeric proteins in which the cytoplasmic tail and transmembrane regions were swapped between the HRSV G and NDV HN proteins were constructed. The chimaeric protein containing the C terminus of the G protein linked to the cytoplasmic tail and transmembrane regions of the HN protein were O-glycosylated normally. Previously, Olmsted *et al* (1989) had demonstrated C terminal truncations of the G protein to be O-glycosylated. SV40 recombinant viruses expressing either 71, 180 or 230 of the N terminal amino acids were found to be O-glycosylated. Furthermore, addition of a 13 amino acid peptide to the C termini of these truncated proteins helped confirm the surface location of these proteins. Expression of these 'chimaeric' proteins from recombinant vaccinia virus followed by immunofluorescence analysis with

antisera raised against the 13 amino acid peptide confirmed the surface location of these proteins, although detection of the 71 N terminus protein was faint. These truncated mutants had previously been used by Vijaya *et al* (1988) to demonstrate the N terminus hydrophobic region to be the transmembrane region. A reporter peptide consisting of repeats of an immunodominant epitope of the circumsporozoite protein of *Plasmodium falciporum* was linked to the C termini of these truncated proteins and used to make recombinant vaccinia virus. A MAb directed against this reporter peptide was used to show cell surface expression of these proteins by immunofluorescence. However no surface expression of the hybrid protein containing the 71 N terminus of HRSV G protein was observed. Lichtenstein *et al* (1991) have demonstrated this hydrophobic region to be the transmembrane region of HRSV. They further demonstrated that this hydrophobic domain consisted of two smaller hydrophobic regions in which only the second hydrophobic region was required for membrane insertion. The first hydrophobic region was found to be non-essential although the possibility of it acting as a transmembrane signal could not be ruled out. Further analysis by Olmsted *et al* (1989) evaluated the immune response of cotton rats to these truncated proteins. Only inoculation with the 230 amino acids truncated protein was capable of inducing an immune response as measured by serum antibody titres and response to HRSV challenge of immunise cotton rats.

Wertz *et al* (1989) expressed the HRSV G protein in a mutant Chinese hamster ovary (CHO) cell line defective in O-glycosylation which was reversible by the addition of galactose and N-acetylgalactosamine (NAG) to the medium. Infection of this cell line with recombinant vaccinia virus expressing the HRSV G protein and varying the composition of the serum free medium with respects to the sugar content enabled the analysis of O-glycosylation of the G protein. In the absence of any sugar and in the presence of tunicamycin only the unglycosylated form (36 kDa) of the protein was observed. In the absence of tunicamycin a protein of 45 kDa representing

only N-glycosylated G protein was observed. The sugars attached to this protein were found to be of the simple type. Addition of galactose to the serum free medium and the absence of tunicamycin resulted in the production of heterodisperse bands greater than 45 kDa representing maturation of the N-linked sugars to the complex form. In the presence of tunicamycin, where all N glycosylation was inhibited, and NAG, but not galactose, O-glycosylated protein with a molecular weight between 67-68 kDa was observed. In the presence of both galactose and NAG but not tunicamycin the fully glycosylated G protein with a molecular weight between 88-90 kDa was observed. The data demonstrates extensive O-glycosylation of the G protein with some contribution by N-linked carbohydrates. Pulse chase experiments showed the 36 kDa unglycosylated form to undergo co-translational N-linked glycosylation which matured to the complex form and then the addition of O-linked carbohydrate. Furthermore surface expression of the G protein measured by indirect immunofluorescence and flow cytometry in the absence of either N- or O-glycosylation was observed. However, cell surface expression of the unglycosylated protein was severely reduced when N- and O-glycosylation were inhibited. These results have corroborated the work of Collins and Mottet (1992) using three inhibitors to identify the glycosylation process of HRSV G protein. In the presence of tunicamycin and carbonylcyanide *m*-chlorophenylhydrazine (CCCP, an inhibitor of protein transport from the ER), or monensin, (an inhibitor of protein transit between the *medial* and *trans* Golgi complex) only the unglycosylated form of the G protein was observed. However, in the presence of tunicamycin and Brefeldin A, (an inhibitor of vesicular transport from the endoplasmic reticulum (ER) which causes disintegration of the medial Golgi complex into either the ER or a ER-Golgi intermediary), the G protein accumulated as a smear with a molecular weight of 68K. Further analysis confirmed this species to be O-glycosylated since it was sensitive to digestion with endo H. Furthermore, this species was not capable of maturing to the complex form since sialic acid residues had been added to the protein. Since such

residues are not normally associated with the G protein, their presence in the species identified above could be a consequence of the action of Brefeldin A upon the Golgi complex (Collins and Mottet, 1992). These results suggest the addition of O-linked carbohydrate occurs within the *trans* Golgi compartment with maturation of the carbohydrate occurring in a subsequent compartment. These results support the work of Lambert (1988) who first demonstrated sensitivity of HRSV G protein to digestion with endo-β-N-acetylgalactosaminidase, O, N or F glycanase. The first two remove O-linked glycans whilst the second two remove N-linked glycans.

Palomo *et al* (1991) investigated the contribution of the G protein carbohydrate content towards immunity. Of a panel of 18 MAb's raised against the HRSV G protein in BALB/C mice inoculated with the Long strain of HRSV (Garcia-Barreno *et al*, 1989), 5 reacted with the unglycosylated form whereas the majority only recognised O-glycosylated G protein that lacked N-glycosylation. Martin-Gallardo *et al* (1993) expressed and purified the unglycosylated HRSV G protein in *Salmonella typhimurium*. Following immunisation of cotton rats with this protein, antisera was shown to contain antibodies that neutralised HRSV *in vitro*. These results suggest a crucial role for carbohydrates for evasion of host immune responses. In the same study by Palomo *et al* (1991) human antisera known to react with the G protein failed to prevent binding of these MAb's to the G protein in competitive ELISA experiments. Garcia-Barreno *et al* (1990) characterised at the molecular level changes within the G protein responsible for generation of HRSV escape mutants against neutralising anti-G MAb (63G; Garcia-Barreno *et al*, 1989). Following a first round of plaque purification isolates were found to contain heterogeneous populations of escape mutants. Following a second round of plaque purification the escape mutants were found to be homogenous. They suggest that following the second round of plaque purification only those that were best adapted survived. Most escape mutants had arisen because of frame shift mutations in which insertion or deletion of A residues altered the coding

region of a third of the C terminus of the G protein. These mutants had a truncation of 25 amino acids at their C terminus. However, one of the mutants had a deletion of a single A residue at position 623 and an insertion of a single A residue at position 648 of the G gene. This resulted in a change of seven amino acids between positions 205 to 211 of the amino acid sequence. Binding of the 63G MAb to synthetic peptides from around this region found amino acids 204 to 209 to be crucial in recognition by the MAb (Garcia-Barreno *et al*, 1992). The authors suggest this domain to be involved in receptor binding since it is located near the four conserved cysteine residues, which had been identified by Johnson *et al* (1987b) as important, and that binding of the MAb 63G interferes with binding of the G protein to the cell surface. These results are made even more significant by the findings of Sullender *et al* (1991) who demonstrated frame shifting to occur in naturally occurring isolates of HRSV subgroup B. Rueda *et al* (1991) generated HRSV escape mutants against five anti-G MAb's. Sequence analysis of these escape mutants revealed the presence of in-frame premature stop codons that shortened the G protein by 11 to 42 amino acids. However, two of the mutants had amino acid substitutions instead of premature stop codons. Sullender *et al* (1991) and Cane *et al* (1991) have identified naturally occurring isolates of HRSV that show extensive genetic diversity in their G proteins. Amino-acid homology within the ectodomain region of the G protein varied between 80%-99% identity between subgroup A viruses (Cane *et al*, 1991) and 88%-97% identity between subgroup B viruses (Sullender *et al*, 1991). However there is only a 53% identity between the A and B subgroup viruses (Johnson *et al*, 1987b). Thus the mechanisms of amino acid substitutions and frame shifting confirms the capacity of the G protein to evolve in response to immunological selection pressures (Sullender *et al*, 1991). The potential effects on virus pathogenicity of these alterations is not known.

Although Ling and Pringle (1989b) identified two forms of the PVM G protein, both were found to be virion associated. Two forms of the HRSV G protein

have also been observed - one associated with the virion and also a soluble form (G_S) secreted into the medium. Hendricks *et al* (1988) characterised the soluble form of the HRSV G protein and have shown that it lacks the first 65 or 74 amino acids covering the transmembrane region. However, Roberts *et al* (1993) have recently described the soluble form of the G protein to be a result of translation from an initiation codon located within (or just beyond) the transmembrane coding region of the HRSV G protein at residue 48, the product of which is subsequently proteolytically cleaved at the N-terminus. Similar result have been observed for the VSV and rabies virus G protein. Chen and Huang (1986) showed the G protein of a ts mutant of VSV to have a mutant G protein that was unstable at the higher temperature, thus allowing access to proteolytic enzymes. The secreted form of the rabies virus G protein was shown to lack the C-terminal membrane anchoring region as a result of proteolytic cleavage during intracellular processing (Morimoto *et al*, 1993).

It is not known if there is an interaction between the HRSV G and M proteins during viral maturation although Nagy *et al* (1991) have demonstrated the HN protein of NDV to play a role in production of viral envelopes. Expression of the NDV HN protein from recombinant baculovirus resulted in the formation of enveloped particles that were indistinguishable from those in NDV infected cells. They could not rule out the possibility of a baculovirus protein substituting for the NDV M protein even though no baculovirus nucleocapsids were observed within these particles. However, this result does suggest a role for the NDV HN protein in assembly and maturation of the virus particle.

1.6.4.4. HRSV GLYCOPROTEINS AND SERUM ANTIBODY RESPONSES :

Taylor *et al* (1984) demonstrated the importance of the F and G proteins in eliciting neutralising antibodies. The N and P proteins were found to be poor antigens. These results were corroborated the findings of Connors *et al* (1991)

who showed that immunisation of mice with recombinant vaccinia virus expressing the P, M, SH, 1C or 1B proteins failed to confer protection following HRSV challenge. Protection against virus challenge was found to significantly decrease in a short period of time following immunisation with recombinant vaccinia virus expressing the N or M2 (22K) protein. However, recombinant vaccinia virus expressing the F or G protein conferred protection against subsequent HRSV challenge. Routledge *et al* (1988) also demonstrated antibody response following immunisation with purified P, M2 (22K), F or G proteins. Only immunisation with the F and G proteins induced production of neutralising antibodies and conferred protection to subsequent HRSV virus challenge. Stott *et al* (1987) used recombinant vaccinia virus expressing the F, G ,N or 1B proteins to investigate the contribution towards immunity of individual proteins. Immunisation of cotton rats with recombinant vaccinia virus expressing the F or G proteins conferred protection to subsequent live virus challenge. However, the protection conferred by immunisation with the F protein was heterologous whereas that by the G protein was homologous. No protection was conferred following immunisation with either the N or 1B proteins. Interestingly, histopathology of the lungs of cotton rats vaccinated with the recombinant vaccinia virus expressing the F, G or N proteins revealed the presence of lesions within the lungs prior to challenge with HRSV. No such lesions were observed following vaccination with the recombinant vaccinia virus expressing the 1B protein. The authors suggest a correlation between these lesions and protection conferred by the various proteins. Similar results demonstrating protection against subsequent virus challenge in cotton rats or mice were observed using either vaccinia virus recombinants expressing the HRSV virus F protein (Wertz *et al*, 1987, Johnson *et al*, 1987a), G protein (Olmsted *et al*, 1986; Elango *et al*, 1986b; Stott *et al*, 1986, Johnson *et al*, 1987a) or affinity chromatography purified F or G protein (Walsh *et al*, 1987).

Garcia-Barreno *et al* (1989) selected escape mutants to an anti-F MAb (47F). These mutants were shown to bind to other anti-F MAb's that bound to a separate antigenic site. They were able to identify four distinct antigenic sites recognised by these MAb's. On the other hand, escape mutants raised against anti-G MAb (63G, described above) failed to react with other anti-G MAb's indicating the existence of structurally overlapping epitopes for the G protein. Wathen *et al* (1989b) expressed a chimaeric protein, termed FG, containing the signal and extracellular regions of the HRSV F protein (amino acids 1 to 489) followed by the extracellular domain of the HRSV subgroup A G protein (amino acids 97 to 279) from a recombinant baculovirus. This protein was secreted into the medium when expressed in insect cells *in vitro* since the chimaeric protein lacked the transmembrane regions of either the HRSV F or G proteins. Immunisation of cotton rats with a crude preparation of this chimaeric protein followed by HRSV challenge demonstrated the induction of neutralising antibodies to this protein (Brideau *et al*, 1989). However the FG chimaeric protein conferred greater protection against infection against subgroup A virus than that conferred by immunisation with a baculovirus secreted form of the F protein. Protection following immunisation with the FG protein was similar to that conferred by the secreted F protein when challenged with a subgroup B virus. These results suggest the G protein contributes significantly towards immunity. These results also highlight the difficulty in generating a vaccine against HRSV. Although the F proteins of the two subgroups of HRSV are fairly well conserved (Collins *et al*, 1984, Elango *et al*, 1985, Baybutt and Pringle, 1987, Johnson and Collins, 1988), there are significant differences between the G protein of not only both subgroups but also between isolates within the same subgroup (Anderson *et al*, 1985, Akerlind *et al*, 1988, Cane *et al*, 1991, Garcia-Barreno *et al*, 1989, Sullender *et al*, 1991).

1.7. AIMS OF THE PROJECT :

Studies of the proteins of HRSV have confirmed the importance of the two major glycoproteins in eliciting an immune response in cotton rats and mice. There is considerable evidence, presented above, demonstrating the ability of the F protein to provide heterologous protection against HRSV infection in experimental animals. However since neutralising antibodies against the F protein have been observed in human sera, the question arises as to how HRSV can re-infect individuals. The diverse variation of the HRSV G protein at the genetic level and the role of glycosylation in evading the immune response suggest a crucial role for the G protein in re-infection. Sequence analysis of PVM genes has demonstrated the similarity to the HRSV genes. PVM offers the opportunity to study a *Pneumovirus* infection in its natural host, the mouse. Two strains of PVM were available, one that is highly pathogenic when inoculated in mice and the other a non-pathogenic strain. The aim of this project was to investigate the molecular basis for the differing pathogenicity with respect to the two major viral glycoproteins by determining the nucleotide sequence of the virus genes. Any differences observed between the two strains may give an insight into the molecular basis of *Pneumovirus* pathogenicity as well as providing comparative information about the genetic structure of pneumoviruses and their relationship to other paramyxoviruses.

CHAPTER 2

MATERIALS AND METHODS

2.1. SOLUTIONS AND BUFFERS :

TE (pH 8.0) : 10mM Tris-HCl (pH 8.0), 0.1mM EDTA.

TNE (pH 8.0): 10mM Tris-HCl (pH 8.0), 100mM NaCl, 1mM EDTA.

PBS (pH 7.4): 136mM NaCl, 2.7mM KCl, 1.3mM KH_2PO_4, 1.4mM Na_2HPO_4.

20X SSC (pH 7.4): 3M NaCl, 300mM Tri-sodium citrate.

10X TAE (pH 8.0): 0.4 M Tris-acetate (pH 7.6), 10 mM Na_2 EDTA (pH 8.0)

10X TBE (pH 8.0): 0.89M Tris-HCl (pH 8.0), 0.89M Boric Acid, 20mM EDTA (pH 8.0).

6X Tracking buffer : 0.25% (w/v) bromophenol blue, 30% (v/v) glycerol.

3X Boiling mix : 15% (v/v) 2-mercaptoethanol, 6% (w/v) SDS, 100mM Tris (pH 6.7), 30% (v/v) glycerol. Bromophenol blue was added till a dark blue colour was obtained.

Phenol : Distilled phenol was melted at 65°C and equilibrated with several changes of equal volumes of TNE.

Phenol/Chloroform : Consisted of a mixture containing equal volumes of phenol and chloroform.

2.2. CELLS :

2.2.1. Bacterial cells used in cloning experiments :

E. coli strain TG2 was the chosen strain for all cloning work except when plasmids carrying tetracycline resistance were used in which case the host was *E. coli* strain HB101. The genotype of both strains is shown below :

E. coli TG2 : supE hsdΔ5 thiΔ(lac-proAB) Δ(srl-recA)306::Tn10(Tet^r)

E. coli HB101 : supE44 hsdS20(r_B^- m_B^-) recA13 ara-14 proA2 lacY1 galK2 rpsL20 xyl-5 mtl-1

2.2.2. Mammalian Cells :

L929: Mouse connective tissue (fibroblast) cell-line used to transiently express the G gene of PVM.

Cos-7: SV40 transformed African Green Monkey fibroblast cell-line used in expression from SV40 based expression vector.

HeLa S3: Human cell-line derived from epithelial carcinoma used in the propagation of vaccinia virus.

BS-C-1: African Green Monkey kidney cells were used in the propagation of PVM and titration of both PVM and vaccinia virus.

HuTK$^-$ 143 : were used in the selection of recombinant Vaccinia virus.

2.3. MEDIA :

2.3.1. Tissue Culture :

In all tissue culture work cells were grown and maintained in GMEM supplemented with non-essential amino acids, 0.2mM glutamine, 100 units/ml of penicillin, 100 µg/ml of streptomycin and FCS (Gibco-BRL or Applied Protein Products Ltd.). The final concentration of FCS depended on the cells and is indicated where appropriate.

2.3.2. Luria-Bertani (LB) medium :

Medium contained the following ingredients per litre :

Tryptone	10 g
NaCl	5 g
Yeast extract	10 g

Medium was sterilised by autoclaving at 121°C for 20 mins. LB agar plates were made with the same ingredients but containing 1.5% (w/v) Bacto-agar and

sterilised as above. Indicator plates for selection of recombinant clones also contained 0.5mM X-Gal and 1mM IPTG.

2.3.3. H-Top medium :

H-top medium was used in transformation of M13 and contained the following per litre :

Tryptone	8 g
NaCl	5 g
Yeast extract	5 g
Bacto-agar	7.5 g

Sterilised by autoclaving for 20 mins at 121°C.

2.4. PREPARATION OF PLASMID DNA :

2.4.1. Small scale plasmid preparation (mini preps) :

The method used was a slightly modified version of that of Serghini *et al* (1989). A culture was grown overnight at 37°C with minimal shaking. 1.5 mls was centrifuged for 5 mins at 15,000g and the pellet resuspended in 100µl of TNE buffer and 100µl of phenol/chloroform followed by centrifugation for a further 5 mins at 15,000g. The top, aqueous, layer was transferred to fresh microcentrifuge tubes to which 35µl of a 10M solution of ammonium acetate and 270µl of cold 100% ethanol had previously been added. The mixture was incubated at 4°C for 20 mins and the nucleic acids were pelleted by centrifugation for 10 mins at 15,000g. The supernatant was removed and traces of ethanol removed by desiccation under vacuum. The pellet of nucleic acids was resuspended in either 30 or 40µl of distilled water since the yields of plasmid DNA obtained were dependent upon plasmid copy number. Aliquots of 10µl were used in restriction enzyme analysis in the presence of 5 µg of RNase A.

2.4.2. Medium scale plasmid preparation (maxi preps) :

Plasmid DNA was prepared using the alkali lysis method of Birnboim and Doly (1979). Unless otherwise stated all steps were carried out at 4°C. Bacteria from a 50 ml overnight grown culture were pelleted for 10 mins at 2,000g. The cell pellet was resuspended in 3 mls of solution I (50mM glucose, 10mM EDTA (pH 8.0) and 25mM Tris-HCl (pH 8.0) solution containing lysozyme at a final concentration of 5 mg/ml). After a 10 min interval 6 mls of solution II (1% (w/v) SDS and 0.2M NaOH). was added. The solutions were mixed by inversion and placed on ice for 10 mins. 4.5 mls of cold 5M potassium acetate (pH 4.8) was added, mixed by inversion and kept on ice for a further 10 mins before centrifugation for 15 mins at 2,000g. The supernatant was removed and mixed with an equal volume of buffered phenol and the two immiscible phases separated by centrifugation for 15 mins at 2,000g. The top, aqueous, layer was removed and mixed with an equal volume of $CHCl_3$ and the two immiscible phases separated by centrifugation as before. The aqueous layer was transferred to fresh tubes and 1/10 volume of a 3M sodium acetate solution (pH 5.5) and 2.5 volumes of cold 100% ethanol added. The mixture was placed on dry ice for 15 mins and the nucleic acids precipitated by centrifugation for 30 mins at 2,000g. The supernatant was removed and the nucleic acids pellet washed with 5 mls of cold 70% ethanol. DNA was recovered by centrifugation for 15 mins at 2,000g. Traces of ethanol were removed by desiccation under vacuum. The nucleic acids were resuspended in 250µl of distilled H_2O.

2.4.3. Large scale plasmid preparation :

The method used is a modified version of the maxi-prep protocol described above. Unless otherwise stated all steps were carried out at 4°C. Briefly, cells from a 500 ml overnight culture were centrifuged at 4,000g for 45 mins, resuspended in 8 mls of solution I and incubated at room temperature for 5 mins.

10 mls of solution II was added, mixed by inversion and the mixture placed on ice for 10 mins. 12 mls of cold 5M potassium acetate (pH 4.8) were added , mixed by inversion and placed on ice for a further 10 mins and centrifugated for 30 mins at 30,000g. The cleared supernatant was removed and mixed with an equal volume of a phenol/chloroform solution and the two immiscible phases separated by centrifugation for 30 mins at 2,000. The aqueous layer was transferred to fresh tubes, 0.6 volumes of isopropanol added and the mixture left at room temperature for 15 mins. DNA was recovered by centrifugation for 30 mins at 2,000g. The alcohol layer was removed and the nucleic acid pellet desiccated under vacuum. The pellet was resuspended in 1.5 mls distilled water and RNase added (1μg/ml final concentration) with incubation at 37°C for 30 mins.

2.4.4. Preparation of M13 single stranded DNA :

5 mls of LB medium inoculated with a single M13 plaque were grown overnight at 37°C with minimal shaking. 1.5 mls of M13 infected *E. coli* cells were centrifuged at 15,000g for 5 mins. 1.2 mls of the supernatant was transferred to fresh tubes to which 250μl of a PEG (33.3mM Polyethylene glycol 6000 (PEG), 2.5M NaCl) solution had been added, mixed by flicking and incubated at room temperature for 15 mins or overnight at 4°C. Precipitated M13 was pelleted by centrifugation at 15,000g for 4 mins and the supernatant discarded. Any remaining supernatant was removed after a further centrifugation for 10 seconds at 15,000g. Pelleted M13 was then resuspended in 120μl of Tris-HCl (pH 8.0) and mixed with 50μl of buffered phenol. The two immiscible phases were the separated by centrifugation at 15,000g for 5 mins. 100μl of the top aqueous layer were mixed with 50μl of $CHCl_3$ and the two phases separated as before. 80μl of the aqueous layer was transferred to fresh microcentrifuge tubes to which 10μl of 3M sodium acetate and 250μl of cold 100% ethanol had been added. The mixture was placed on dry ice for 10 mins and the single-stranded DNA pelleted by centrifugation at 15,000g for 10 mins.

2.4.5. CsCl Gradient DNA purification :

DNA prepared using the large scale plasmid prep was diluted to a final volume of 3.8 mls with distilled water and 4.2 g of CsCl and 50µl of a 10 mg/ml solution of ethidium bromide were added. The mixture was transferred to Beckman heat seal tubes (5 ml) and topped up with a solution of 3.8 ml water, 4.2 g CsCl and 50µl of 10 mg/ml solution of ethidium bromide. The tubes were sealed and centrifuged for at least 18 hours at 45,000 rpm at 20°C in a Beckman VTi65 rotor.

The visible band of DNA was excised from the CsCl gradient by side puncture and the ethidium bromide removed by extraction 6 times with equal volumes of amyl alcohol. Plasmid DNA was dialysed overnight against at least 3 changes of 5 litres of a 0.1X SSC in order to remove the CsCl. The DNA was precipitated by addition of 1/10 volume of 3M sodium acetate (pH 5.5) and 2.5 volumes of 100% ethanol, stored at either -70°C for 4 hrs or overnight at -20°C and then centrifuged at 2,000g at 4°C for 30 mins. Traces of alcohol were removed by desiccation under vacuum and the total concentration of plasmid DNA estimated by optical density (O.D.) reading at 260 nm. 1 O.D. unit was estimated as 40 mg/ml of DNA.

2.5. TRANSFORMATIONS :

2.5.1. Transformation of E. coli :

2.5.1.1. PREPARATION OF COMPETENT *E. COLI* :

The method used was based on that of Cohen *et al* (1972). 70 mls of LB medium, inoculated with 1 ml of stationary phase *E. coli* were grown at 37°C with minimal shaking to an O.D. reading at 650 nm of approximately 0.4 and subsequently incubated on ice for at least 30 mins. 50 mls of culture was centrifuged at 2000g at 4°C for 5 mins, resuspended in 25 mls of a ice-cold solution of 0.1M $MgSO_4$ and pelleted once again at 2000g at 4°C for 5 mins. The bacteria were resuspended in

2.5 mls of an ice-cold solution of 0.1M $CaCl_2$ and incubated on ice for 1 hour before beginning the transformation procedure.

2.5.1.2. TRANSFORMATION OF COMPETENT *E. COLI* WITH PLASMID DNA :

DNA was added to an aliquot (100µl) of competent bacteria and incubated on ice for 30 mins, transferred to 42°C for 2 mins, replaced on ice for a further 5 mins and then left at room temperature for 5 mins. 1 ml of sterile LB medium was added and the mixture incubated at 37°C with minimal shaking for 40 mins. 0.33 ml aliquots were then plated onto LB agar plates containing the appropriate antibiotic selection and the plates incubated overnight at 37°C.

2.5.1.3. TRANSFECTION OF COMPETENT CELLS WITH M13 :

The protocol was similar to that described above except that after the 42°C incubation, bacteria were kept on ice for 5 mins and then added to a mixture of 3 mls of molten H-Top agar containing 75mM X-Gal, 150mM IPTG. 80µl of an *E. coli* culture that had been grown overnight, diluted 1/100 with LB medium and incubated at 37°C with minimal shaking for 1 hour was added immediately prior to plating out. The mixture was poured onto LB plates and the H-top allowed to set before overnight incubation at 37°C.

2.5.2. Transfection of mammalian cells :

2.5.2.1. CALCIUM PHOSPHATE CO-PRECIPITATION :

DNA was introduced into mammalian cells using a modified protocol of Chen and Okayama (1987). Monolayer of cells were grown to 85-90% confluence in 5 cm plates and maintained in GMEM supplemented with FCS (5% final concentration) at 37°C in a 5% CO_2 atmosphere. 20 µg of CsCl gradient purified DNA was mixed with 250µl of a BBS solution (pH 6.96) [50mM BES (N,N-bis (2-hydroxyethyl)-2-aminoethanesulfonic acid), 280mM NaCl, 1.5mM Na_2HPO_4]. 125µl of a 2M $CaCl_2$ solution were added dropwise to the mixture whilst N_2 was gently

bubbled through to ensure mixing. The resultant mixture was kept at room temperature for 20 to 30 mins and then added dropwise the monolayer of cells. The cells were maintained at 37°C, in an atmosphere of 3% CO_2 for between 6 and 8 hours, washed twice with 2 mls of sterile PBS and osmotically shocked with 2 mls of a 15% glycerol in PBS solution for 2 mins at room temperature. Cells were washed twice with PBS, once with GMEM supplemented with FCS to a final concentration of 5% and then maintained in 4 mls of GMEM supplemented with FCS to a final concentration of 5% at 37°C in a 5% CO_2 atmosphere. Transfected cells were harvested after a further 48 hours.

2.5.2.2. TRANSFECTACE™ REAGENT (GIBCO-BRL) :

5 µg of plasmid DNA was diluted with 100µl of Opti-MEM® (Gibco-BRL) reduced medium and added to 12µl of TransfectACE™ reagent (diluted in 100µl Opti-MEM® reduced medium) in a polystyrene tube. The TransfectACE™ reagent-DNA complex was left at room temperature for 15 mins. Opti-MEM® reduced medium was added to a final volume of 1 ml and the total mixture added to cells that had been grown to 85-90% confluency at 37°C in 5% CO_2. After a 6 hr incubation at 37°C, 5% CO_2, 1 ml of GMEM supplemented with FCS at a final concentration of 20% medium. Cells were incubated at 37°C, 5% CO_2 for 48 hours before harvesting.

2.6. ENZYMES :

2.6.1. Restriction endonucleases :

DNA to be analysed by restriction endonuclease digestion was prepared either by the mini-, maxi- or large scale DNA preparation methods described above. All digests were carried out in the manufacturer's recommended buffers in the presence of 1-2 units of enzyme and 0.1 mg/ml BSA. For DNA prepared by the mini- and maxi preparation methods, 100 µg/ml RNase A was also added. All digests were carried out

at 37°C for between 2-6 hours. Restriction enzymes were supplied by either Amersham, Gibco-BRL or Pharmacia.

2.6.2. De-phosphorylation of linearised vector DNA :

To minimise vector re-ligation the 5' terminal phosphate of linearised DNA was removed. 2-5 units of calf intestinal phosphatase (CIAP) was added during the last hour of restriction endonuclease digestion of vector DNA.

2.6.3. Klenow DNA polymerase :

1mM each of dATP, dTTP, dGTP, dCTP and 1-2 units of Klenow DNA polymerase were added to the restriction endonuclease digest reactions and incubated for 30 mins at 37°C in order to end-fill digested DNA.

2.6.4. T4 DNA polymerase :

T4 DNA polymerase possesses an equivalent 5'→3' polymerase but more active 3→5' exonuclease activity than the large fragment of *E. coli* DNA polymerase I (Klenow Fragment). In the presence of excess deoxynucleotides, the amount of 3'→5' exonuclease activity is reduced considerably when compared to 5'→3' polymerase activity. Additionally the 3'→5' exonuclease activity is only functional on ssDNA or on dsDNA with 3' overhangs but not on 5' overhangs (3' recess). Thus, T4 DNA polymerase was chosen over DNA polymerase I (Klenow fragment) for blunting of 5' and 3' overhangs of dsDNA following digestion with restriction enzymes.

Restriction endonuclease treated DNA was first phenol/chloroform extracted and then ethanol precipitated as described in the mini- DNA preparation protocol. Linearised DNA was blunted in the presence of 1mM each of dATP, dTTP, dGTP, dCTP, 33mM Tris-HCl (pH 7.5), 66mM potassium acetate, 10mM magnesium acetate, 0.5mM DTT, 0.1 µg BSA and 1-2 units of T4 DNA polymerase and the reaction incubated at 37°C for 45 mins. Blunted DNA was then either purified from LMP agarose gels or phenol/chloroform extracted and ethanol precipitated.

2.6.5. T4 DNA Ligase :

Ligation of DNA fragments was carried out overnight at 16°C in the presence of 66mM Tris-HCl (pH 7.6), 5mM $MgCl_2$, 5mM DTT, 0.1 µg BSA, 1mM hexaminecobalt chloride, 1mM ATP, 0.5mM spermidine HCl and 1-2 units T4 DNA ligase. Typically, the vector DNA to insert DNA ratio varied between 1:2 and 1:5 in a final reaction volume of 20-25µl. DNA fragments purified from LMP agarose gels were heated to 65°C for 10 mins and the ligation carried out in a final volume of between 60-100µl, where the final volume was in at least 3 fold excess of the volume of DNA fragment added.

2.6.6. Sequenase® sequencing reactions :

Sequencing reaction of ssDNA (M13 DNA) was performed using the Sequenase® (version 2.0) sequencing kit (U.S. Biochemical Corporation). 7 µl of ssDNA (1 µg) was annealed to 1 µl primer (0.5 pmol) after the addition of 2 µl 5X reaction buffer (200 mM Tris-HCl (pH 7.5), 100 mM MgCl2, 250 mM NaCl) by heating to 65°C for 2 mins and then allowed to cool to room temperature over a period of 30 mins. Primer extension reactions were performed at room temperature for 4 mins after the addition of 1 µl 0.1M DTT, 2 µl labelling mix (7.5 µM each of dCTP, dGTP and dTTP: diluted 5 fold with distilled water), 0.5 µl α-35S dATP and 2 µl of Sequenase version 2.0 enzyme diluted 8-fold in enzyme dilution buffer (10mM Tris-HCl (pH 7.5), 5mM DTT, 0.5 mg/ml BSA). 3.5 µl aliquots were divided between four tubes containing 80 µM each of dCTP, dATP, dGTP, dTTP and 8 µM of ddCTP or ddATP or ddGTP or ddTTP. These termination reactions were carried out at 37-42°C for 4 mins. The reactions were stopped by the addition of 4 µl of stop solution (95% formamide, 20mM EDTA, 0.05% (w/v) bromophenol blue and 0.05% xylene cyanol FF). The sequence reactions were electrophoresed as described in section 2.7.1.4.

2.6.7. Exonuclease *Bal31* :

Exonuclease *Bal31* possesses a 5'→3' single stranded endonuclease and a 3'→5' exonuclease activity. Its was therefore used to create sequential deletions from the ends of linearised dsDNA fragments. Typically, the reaction was carried out in the presence of 0.6 M NaCl, 12mM $CACl_2$, 12mM $MgCl_2$, 20mM Tris-HCl (pH 8.0), 0.2mM EDTA (pH 8.0) and 1-2 units of *Bal31* at 30°C. Aliquots were taken out at various intervals and the reaction was stopped by addition of EDTA to a final concentration of 3mM. Since the 3'→5' exonuclease activity is greater than the 5'→3' endonuclease activity the ends of *Bal31* treated DNA fragments possess a 5' overhang. Thus, all *Bal31* digested fragments were phenol/chloroform extracted, ethanol precipitated followed by end-blunting of the fragments with T4 DNA polymerase as described.

2.6.8. Isolation of PVM infected mRNA :

BS-C-1 cells were infected at a multiplicity of infection of 1 and the inoculum replaced with GMEM supplemented with FCS to a final concentration of 2%. The cells were incubated at 37°C for 24 hours and the medium replaced with GMEM containing FCS (2% final concentration) and actinomycin D (2.5 μg/ml) and incubated at 37°C for 48 hours. PVM infected cells were harvested and pelleted at 4000 g for 5 mins at 4°C. The RNA extraction was carried out using a method based on that of Kumar and Lindberg (1972). Pelleted cells were resuspended in isotonic lysis buffer (150 mM NaCl, 1.5 mM Mg_2Cl, 10 mM Tris (pH 7.8), 0.65% NP40) and placed on ice for 10 mins. The cell debris was pelleted for 5 mins at 4000 g, 4°C and the supernatant extracted with 10 mls of extraction mix (containing equal volumes of phenol/chloroform (phenol was equilibrated with 150 mM NaCl, 1.0 mM EDTA, 10 mM Tris (pH 7.8)) and phenol extraction buffer (7.0 M urea, 350 mM NaCl, 10 mM EDTA, 1% SDS, 10 mM Tris (pH 7.8)) The aqueous phase was re-extracted twice with phenol\chloroform and RNA precipitated overnight at -20°C from the final

aqueous phase following the addition of 2.5 volumes of ethanol. The precipitate was resuspended in water to a final concentration of 1 mg\ml.

2.6.9. AMV reverse transcriptase : cDNA synthesis from cellular RNA :

The first strand of cDNA was synthesised from cellular RNA of PVM-infected cells using avian myeloblastosis virus (AMV) reverse transcriptase (Life Sciences, Inc.). Reverse transcription was carried out in the presence of 0.5mM each of dATP, dCTP, dGTP, dTTP, 50mM Tris-HCl (pH 8.3), 6mM $MgCl_2$, 40mM KCl, 3µg of poly d(T) (pdT; Pharmacia) and 10 units AMV reverse transcriptase. The reaction mixture was incubated at 42°C for 45 mins, phenol/chloroform extracted and ethanol precipitated. cDNA was resuspended in 30µl of distilled water. 3µl aliquots were used to amplify DNA fragments using the polymerase chain reaction.

2.6.10. In vitro transcription :

T7 RNA polymerase was used to synthesise RNA *in vitro*. Approximately 0.1-0.5µg of linearised DNA was treated with 25 units of the appropriate RNA polymerase in the presence of 1mM ATP, GTP, UTP, CTP, 4mM Tris-HCl (pH 8.0), 8mM $MgCl_2$, 2mM spermidine-$(HCl)_3$, 25mM NaCl, 5mM DTT. The reaction was incubated at 30°C for 1-2 hours. *In vitro* transcribed RNA was phenol/chloroform extracted, ethanol precipitated and resuspended in 20-30µl of distilled water. Where necessary the RNA was stored at -20°C as an aqueous solution.

2.6.11. In vitro translation :

In vitro transcribed RNA was used to synthesise proteins *in vitro* using rabbit reticulocyte lysate (Amersham). The reaction involved incubation of 1µl of *in vitro* transcribed RNA with 8µl of rabbit reticulocyte lysate and 10 µCi ^{35}S-methionine at 30°C for 1 hour.

Alternatively *in vitro* transcribed RNA was translated in the presence of wheat germ extract (gift from Dr. R. Wales, Dept. Biological Sciences, University of

Warwick). The reaction involved incubation of 1µl of *in vitro* transcribed RNA with 10 µCi ^{35}S-methionine, 3.4µl distilled water, 3.8µl wheat germ extract, 2.4µl energy mix (15mM HEPES (pH 7.6), 2 mM ATP, 4mM creatine phosphate, 1.2 mM DTT, 0.004 mM GTP, 0.004mM amino acids, 2.4 mM potassium acetate, 0.2 mg/ml creatine PK, 0.15 mM spermidine (pH 7.0)). The reaction was incubated for 1 hour at 27°C.

3X reducing boiling mix was added and the reaction electrophoresed on a polyacrylamide gel.

2.6.12. Indirect immunofluorescence :

Transfected cells were scraped into the medium using a sterile rubber policeman and pelleted for 10 mins at 2,000g. The supernatant was removed and the cells were resuspended in 0.5 to 1 ml of GMEM supplemented with FCS to a final concentration of 5%. Aliquots were transferred onto 8-well slides and incubated overnight at 37°C.

Cells were washed twice with PBS (5 mins each) and permeabilised by immersion in cold acetone for 2 x 10 mins. Fixed cells were then either stored at -20°C or used straight away. Cells were washed three times with PBS and incubated PVM anti-G (19/1/C9) monoclonal antibody, diluted 1:100 in PBS, in a moist chamber at 37°C for 1 hour. Cells were washed three times with PBS and then either incubated with 1:400 dilutions of goat anti-mouse FITC or biotinylated sheep anti-mouse antibody in PBS for 1 hour in a moist chamber at 37°C. Cells treated with the biotinylated goat anti-mouse antibody were then washed three times with PBS and incubated with 1:400 dilution of streptavidin-fluorosceine conjugate or streptavidin-Texas red in PBS in a moist chamber at 37°C for 20 mins. Finally, the cells were washed three times with PBS and mountant (Citifluor, Amersham or 80% glycerol in PBS) added.

Fluorescence or Texas-red treated cells were observed under a Zeiss UV microscope and photographs taken with a Kodak ektachrome 160 tungsten film.

2.7. GEL ELECTROPHORESIS :

2.7.1. DNA Gels :

2.7.1.1. AGAROSE GELS :

Restriction endonuclease digested DNA was visualised with UV illumination after electrophoresis on either a 1% (w/v) or 2% (w/v) agarose gel depending upon the size of DNA fragments. Agarose was dissolved in TBE or TAE containing 0.05 µg/ml of ethidium bromide and cast in either a 6 x 8, 8 x 10.5, 11.5 x 14 or 20 x 25 cm horizontal tray. 6X tracking dye buffer was added to samples and the gel electrophoresed at between 100 and 180 volts in the presence of TBE buffer containing 0.05 µg/ml ethidium bromide.

2.7.1.2. LOW MELTING POINT AGAROSE GELS :

1% (w/v) low melting point agarose gels (LMP) were used to purify restriction endonuclease digests of fragments that were subsequently to be used in ligation reactions. Gels were cast and electrophoresed under the same conditions as ordinary agarose gels at a constant 90 volts. DNA fragments were excised from the gel, transferred to fresh Eppendorf tubes and then used in ligations.

2.7.1.3. GENECLEAN™ DNA PURIFICATION :

DNA bands were purified from 1% TAE agarose gels using the GeneClean™ kit (Bio 101 Inc., La Jolla, California, USA) and purified following the manufacturer's instructions. Briefly, excised DNA bands were heated to 55°C for 5 mins following dilution with 3 volumes of NaI stock solution and mixed with Glassmilk® at a ratio of 1µl of Glassmilk® for 1µg of DNA excised. The DNA/Glassmilk mixture was incubated at room temperature for 5 mins, washed

3 times with New Wash solution and the DNA eluted into water after incubation for 3 mins at 55°C.

2.7.1.4. SEQUENCING GELS :

Sequencing reactions were electrophoresed on a 5.7% (w/v) acrylamide, 0.3% (w/v) bis-acrylamide containing TBE and 7M urea. Ammonium persulphate was added to a final concentration of 0.05% (w/v) and 55µl of TEMED was also added per 120 mls of mixture. A 31 x 38.5 cm gel was cast horizontally between two glass plates separated by a wedge spacer generating a gel ranging from 0.4mm to 1.2mm in width. The gel was electrophoresed in TBE running buffer at a constant 80 watts. After electrophoresis the gel was fixed by soaking in 10% (v/v) acetic acid for at least 1 hour. Gels were dried on a Hoefer Scientific Instruments Drygel Sr gel dryer at 80°C for 2 hours and autoradiographed.

2.7.2. Protein gels :

2.7.2.1. POLYACRYLAMIDE GEL ELECTROPHORESIS (PAGE) :

Protein samples to be analysed were electrophoresed on a 10% (w/v) polyacrylamide running gel with a 6% (w/v) stacking gel (Laemmli, 1970).

The 10% acrylamide running gel consisted of 9.67% (w/v) acrylamide, 0.33% (w/v) bis-acrylamide, 0.4mM Tris-HCl (pH 8.8), 3.5mM SDS, 4.4mM ammonium persulphate and 20µl of TEMED per 50 mls of gel. The mixture was cast between two glass plates and 2 mls of butanol added to the top of the gel. The mixture was allowed to polymerise for not less than 2 hrs. The butanol layer was removed and the top of the running gel washed several times with distilled water and dried using 3M blotting paper. The 6% (w/v) acrylamide stacking gel consisted of 5.8% (w/v) acrylamide, 0.2% (w/v) bis-acrylamide, 0.125mM Tris-HCl (pH 6.8), 3.5mM SDS, 0.9mM ammonium persulphate and 15µl of TEMED per 20 mls of gel. The stacker was cast on top of the running gel and the comb gently inserted.

Polyacrylamide gel electrophoresis was carried out in the presence of a running buffer containing 0.18M glycine, 25mM Tris, 3.5mM SDS. Samples were initially run into the stacker gel for 30 mins at 35 mAmps and then run overnight at between 8-10 mAmps.

Gels were soaked in a solution of 10% (v/v) acetic acid and 40% (v/v) methanol for at least 1 hour and dried on a Hoefer Scientific Instruments Drygel Sr gel dryer at 80°C for 2 hours prior autoradiography.

2.7.2.2. FLUOROGRAPHY OF PROTEIN GELS :

Gels were fixed in 40% (v/v) methanol, 10% (v/v) acetic acid, 5% (v/v) glycerol overnight followed by Amplify (Amersham, UK) for 30 mins and dried for 2 hours at 80°C. Gels were then exposed to pre-flashed Fuji X-ray film and kept at -70°C .

2.8. AMPLIFICATION OF DNA FRAGMENTS BY THE POLYMERASE CHAIN REACTION:

2.8.1. Amplification using Vent™ DNA Polymerase :

Amplification was achieved by 40 cycles of DNA denaturation at 94°C for 1.5 mins, oligonucleotide annealing at 55°C for 1.5 mins and extension at 74°C for between 2 to 5 mins depending upon size of fragment to be amplified, in the presence of 10mM KCl, 10mM $(NH4)_2SO_4$, 20mM Tris-HCl (pH 8.8), 2mM $MgSO_4$, 0.1% Triton X-100, 100 µg/ml BSA, 0.5mM each of dATP, dTTP, dCTP, dGTP and 1 unit of Vent™ DNA polymerase (New England Biolabs, MA, USA).

2.8.2. Amplification using Taq™ DNA Polymerase :

Amplification by PCR was achieved by 40 cycles of; denaturation at 95°C for 1.5 mins followed by oligonucleotide annealing and extension at 65°C for 3 mins. Amplification was performed in 3 µg/ml gelatin, 0.5mM each of dATP, dCTP,

dGTP, dTTP, 0.4mM DTT, 4mM $MgCl_2$, 30mM KCl, 50mM Tris-HCl (pH 8.3) and 10 units/ml of Taq™ DNA polymerase (Amplitaq, Cetus).

2.9. TISSUE CULTURE WORK :

2.9.1. Growth of cells :

Cells were grown in 25 cm tissue culture flasks and passaged at 4-5 day intervals. Cells were rinsed with a versene/trypsin mix and then with 0.25% (w/v) trypsin solution. A small amount was retained and the cells incubated at 37°C until all were detached. They were then resuspended in either GMEM supplemented with FCS to a final concentration of 5% or 10%, seeded at a ratio of 1:4 and maintained at 37°C.

Unless otherwise stated all cells were maintained at 37°C in a 5% CO_2 atmosphere.

2.9.2. Pneumonia Virus of Mice (PVM) :

PVM strain 15 was originally obtained from the American Type Culture Collection (Ling and Pringle, 1989a). PVM strain J3666 was obtained from Dr. D. Harter of the Rockefeller Institute, New York. Both strains were propagated in BS-C-I cells.

2.9.2.1. PREPARATION OF MURINE ERYTHROCYTES :

Mice were killed in a CO_2 saturated environment. Blood was obtained from the thoracic cavity after opening the chest and cutting the aorta. Cells were washed twice with PBS and resuspended in 5 mls of PBS and cell concentration was determined using a Neubauer haemocytometer.

2.9.2.2. HAEMAGGLUTINATION ASSAY (HA) :

Serial two-fold dilutions of PVM strain 15 and J3666 (in PBS) were prepared in 96-well round bottom plates in 100 µl. To each well, 3 x 10^4 mouse red blood cells were added to give a final volume of 200µl. Plates were left at room

temperature for 1-2 hrs and the HA titre determined from the lowest virus dilution in which no agglutination was observed.

2.9.2.3. HAEMAGGLUTINATION INHIBITION ASSAY (HI) :

PVM anti-G monoclonal antibody (19/1/C9) was initially diluted 1 in 10 and then serially diluted two-fold (in PBS) in 96-well round bottom plates. 5 HA units of virus was added together with 3 x 10^4 mouse erythrocytes in a final volume of 200 µl. Plates were left at room temperature for 1-2 hrs. HI titre was determined from the highest antibody dilution that inhibited haemagglutination.

2.9.2.4. TITRATION OF PVM STRAINS 15 AND J3666 :

BS-C-1 cells were grown to confluence as monolayers in either 6-well or 12-well tissue culture dishes. The medium was aspirated and 10-fold dilutions of PVM strain 15 or J3666 (in PBS) were added to separate wells. The cells were incubated for 2 hours at 33°C after which the inoculum was replaced with a 1% (w/v) agar overlay containing GMEM supplemented with FCS to a final concentration of 2%. The agar overlay was allowed to set and the cells incubated at 33°C. When plaques were visible, 2% glutaraldehyde was added and the cells left for at least 5 hours to overnight at room temperature. The agar overlay was removed carefully and the cells were stained with 0.075% (w/v) crystal violet solution (in PBS) for 10 mins. Excess stain was removed by rinsing each well under running tap water. The plates were allowed to dry and the plaques counted.

2.9.2.5. RADIOACTIVE LABELLING OF VIRUS INFECTED CELLS :

BS-C-1 cells were grown to confluence as monolayers in 6-well tissue culture dishes. The medium was aspirated and the cells washed once with sterile PBS. PVM strain 15 or strain J3666 was added to a multiplicity of infection of between 1 and 2 pfu/cell and the cells incubated at 33°C for 2 hours. The inoculum was replaced with GMEM supplemented with FCS to a final concentration of 2% and

incubated at 33°C. When CPE was observed the medium was removed, the monolayer washed once with sterile PBS, and serum free GMEM added with a further incubation at 33°C for 3 hours. A fresh aliquot of serum free GMEM containing either 100 μCi/ml ^{35}S L- Cysteine or 50 μCi/ml of ^{3}H-glucosamine hydrochloride was added and the cells were incubated for 18-24 hours at 33°C.

For PVM infected cells treated with either tunicamycin or monensin, serum free GMEM was supplemented with either 2.5 μg/ml or 0.8μM of tunicamycin or monensin respectively at both the pre-labelling and labelling periods described above.

2.9.2.6. RADIOIMMUNOPRECIPITATIONS (RIP) :

PVM infected, labelled cells were scraped into the medium and transferred to microcentrifuge tubes. Cells were pelleted by centrifugation at 15,000g for 10 mins. The supernatants were transferred to fresh tubes and incubated at 4°C overnight following the addition of PEG to a final concentration of 6% (w/v) Virus was precipitated by centrifugation at 15,000g for 10 mins. The pelleted cells and virus were washed twice in PBS containing 1mM PMSF and resuspended in 100μl of high salt RIP buffer (1% (v/v) Triton X-100, 150mM NaCl, 600mM KCl, 0.5mM $MgCl_2$, 10mM Tris-HCl (pH 7.4)) containing 1mM PMSF and placed on ice for 1 hour prior to storing at -20°C.

Immunoprecipitin (Gibco-BRL, UK) was prepared according to the manufacturer's recommendation to reduce sloughing of formalin fixed *Staphylococcus aureus* cells during the immunoprecipitation procedure. *Staphylococcus aureus* cells were pelleted by centrifugation at 15,000g for 10 mins, resuspended in equal volume of 10% (w/v) 2-mercaptoethanol, 3% (w/v) SDS in PBS and heated to 95°C for 30 mins. Cells were pelleted at 15,000g for 10 mins and resuspended in equal volume of RIP buffer.

50µl of immunoprecipitin (in RIP buffer) and 10µl of FCS were added to the cell lysate and virus preparations and incubated on ice for 30 mins. *Staphylococcus aureus* cells were pelleted for 10 mins at 15,000g and the supernatants transferred to fresh microcentrifuge tubes. 10µl of PVM anti-G monoclonal antibody (19/1/C9- Ling and Pringle, 1989b) was added and the mixture left at 4°C overnight. 80µl of immunoprecipitin was added and the mixture was placed on ice for 30 mins before pelleting the *Staphylococcus aureus* cells at 15,000g for 10 mins. The supernatants were discarded and the cells washed three times with 1 ml of RIP wash buffer (0.5M LiCl, 100mM Tris-HCl (pH 8.9)), containing 1mM PMSF. Finally the cells were resuspended in 50µl of boiling mix and stored at -20°C. Prior to electrophoresis, cells were boiled for 5 mins.

2.9.3. Vaccinia virus :

Wild type vaccinia virus, strain Wr, was obtained from Professor M. A. McCrae, Department of Biological Sciences, University of Warwick. Recombinant vaccinia virus, vTF7-3 (Fuerst *et al*, 1986), that expresses the T7 RNA polymerase was kindly provided by Dr. R. Elliott of the Glasgow Institute of Virology, Glasgow, UK., with the permission of Dr. B. Moss of the National Institute of Allergy and Infectious diseases, Bethesda, Maryland, USA.

2.9.3.1. PREPARATION OF WILD TYPE VACCINIA VIRUS STOCK :

HeLa cells were grown to confluence as monolayers in 175 cm tissue culture flasks. The medium was removed and wild type vaccinia virus added to a multiplicity of infection of 0.01 pfu/cell. Following incubation at 37°C for 2 hours, the inoculum was replaced with GMEM supplemented with FCS to a final concentration of 2% and the cells incubated at 37°C. After 3 days, when extensive CPE was visible, the flasks were kept at -70°C till the medium had frozen. The medium was thawed by incubation at 37°C and this process of freeze/thaw repeated twice. Cell debris and

virus were pelleted by centrifugation at 20,000 rpm at 4°C for 2.5 hours, resuspended in PBS and sonicated for 4 mins in a sonicating bath. Cell debris was pelleted by centrifugation at 4000g at 4°C for 10 mins and the supernatant stored at -70°C.

2.9.3.2. TITRATION OF VACCINIA VIRUS :

BS-C-1 cells were used for titration of vaccinia virus using the method described in section 2.9.2.4, except virus adsorption and growth was at 37°C

2.9.3.3. RECOMBINANT VACCINIA VIRUS :

85-90% confluent monolayers of HeLa cells grown in a 5 cm tissue culture dish were infected with wild type vaccinia virus at a multiplicity of infection of 0.01 pfu/cell. The virus was allowed to adsorb for 2 hours at 37°C. Infected cells were subsequently transfected with CsCl purified DNA of pSC11 constructs containing the G genes of PVM strains 15 and J3666 using the Gibco-BRL TransfectACE protocol described in section 2.5.2.2.

Transfected cells were scraped into the medium, transferred to plastic universal tubes and freeze/thaw treated as described above. Cell debris was pelleted by centrifugation at 2000g at 4°C for 10 mins. Recombinant vaccinia virus was selected by titration of the supernatant.

2.9.3.4. SELECTION OF RECOMBINANT VACCINIA VIRUS :

HuTK$^-$ 143 cells, which lack thymidine kinase activity, were grown to confluence as monolayers. The medium was replaced with supernatant containing recombinant vaccinia virus (prepared above) and the cells incubated for 2 hours at 37°C. The inoculum was replaced with 1% (w/v) agar overlay containing GMEM supplemented with FCS (final concentration of 2%) and 25 µg/ml BrdU. The agar was allowed to set and the cells then incubated at 37°C. After 3-4 days 2 mls of agar overlay containing serum free GMEM and 20 µg/ml of X-gal was added, the agar allowed to set and the cells incubated at 37°C, 5% CO_2. Recombinant vaccinia virus

plaques were thus distinguished from spontaneous revertants by their blue appearance. A sterile Pasteur pipette was inserted into the agar overlay and the blue plaques transferred into 500µl GMEM supplemented with FCS (final concentration of 2%) medium. They then underwent three freeze/thaw cycles and amplified HuTK$^-$ 143 cells.

CHAPTER 3

NUCLEOTIDE SEQUENCE OF PVM FUSION GENES

3.1. INTRODUCTION :

Viral infection of the *Paramyxoviridae* family of viruses is initiated by the concerted action of two major viral surface glycoproteins. The first, designated G, is responsible for the attachment of the virus particle to the host cell receptor and is referred to as the attachment protein. This protein varies in its general characteristics between the morbillivirus, paramyxovirus and pneumoviruses. The second glycoprotein (designated F) is responsible for fusion of viral and cell membrane, allowing entry of the viral genome into the cell, and fusion of infected cell and adjacent cell allowing the spread of the virus (Pringle 1987, Morrison 1988).

Chambers *et al* (1992) have recently published the nucleotide and predicted amino acid sequence of the F gene of a non-pathogenic strain of Pneumonia virus of mice (PVM, strain 15), that has been passaged entirely in tissue culture. They found little (10%) amino acid homology with the F proteins of the morbillivirus and paramyxovirus and only a 40% homology with the F proteins of human respiratory syncytial virus (HRSV) and turkey rhinotracheitis virus (TRTV), the other members of the *Pneumovirus* genus. However the structural features of the protein were consistent with those of the other members of the *Paramyxoviridae* family as described in chapter 1. Another strain of PVM, J3666, was obtained from Dr. D. Harter of the Rockefeller Institute, New York which has been passaged entirely in mice and is highly pathogenic.

Work on virulent and avirulent strains of Newcastle disease virus (NDV) has shown the pathogenicity of each strain to be related to the susceptibility of the F_0 polypeptide to cleavage by certain proteases (Nagai *et al*, 1976, Nagai and Klenk, 1977). It was anticipated that sequence analysis of this gene of PVM strain J3666 might reveal amino acid changes at or in close proximity to the cleavage site of the F_0 polypeptide. Sequence analysis of the F gene of virulent and avirulent strains of NDV have attributed changes at the cleavage site of F_0 to the relative pathogenicity of

the virus. Early tissue culture work by Nagai *et al* (1976) and Nagai and Klenk (1977) showed a correlation between the virulence and susceptibility to cleavage of the F_0 polypeptide. Sequence analysis of the F genes of the different strains of NDV by Glickman *et al* (1988) and Toyoda *et al* (1987) found that cleavage occurred between the amino acids Arg and Phe in all virulent strains of NDV whereas Phe had been substituted by Leu in all avirulent strains . Thus the amino terminus of the F_1 subunit of virulent strains of NDV began with a Phe residue whereas that of the avirulent strains began with Leu. Additionally, in the virulent strains of NDV the putative cleavage site was located within a hexapeptide consisting of four (or five) basic amino acids with an intervening Gln residue (Arg-Arg-Gln-Arg/Lys-Arg-Phe) whereas in the avirulent strains the first and fourth residues were replaced with the neutral amino acid, Gly.

Recently, Morrison *et al* (1993) demonstrated that altering the cleavage site of a virulent strain of NDV (Arg-Phe) to Arg-Gly had no effect on cleavage of the F_0 polypeptide and that the fusion activity was equivalent to that of the wild type. However, when the cleavage site was changed to Arg-Leu, the F_0 polypeptide was no longer cleaved and no fusion activity was observed. The fusion function of this polypeptide was regained after digestion with trypsin. The authors suggested several explanations for these observations of which the most plausible is that the presence of Leu alters the conformation of the F_0 polypeptide and makes the cleavage site inaccessible to proteases during maturation. These findings are somewhat complicated by the work of Horvath and Lamb (1992) on SV5 in which Phe had been substituted with Leu but no effect on either fusion or cleavage was observed. Morrison *et al* (1993) suggested several reasons for this discrepancy. For example, the substitution of Phe by Leu may not alter the 3-D conformation around the fusion domain, or the F_0 polypeptide of SV5 could be cleaved by a different protease to that of NDV.

The F protein of Sendai virus, responsible for respiratory tract infections of rodents and responsible for fatal pneumonia in laboratory mice (Ishida and Homma, 1978), is unusual in comparison to that of other members of *Paramyxoviridae* whose F_0 polypeptide cleavage occurs intracellularly during maturation. Sendai virus F_0 polypeptide is cleaved extracellularly by some proteases in order to activate the fusion activity (Muramatsu and Homma, 1980, Tashiro and Homma, 1983) and Sendai virus host range and cell/organ tropism is determined by the presence of these proteases. Sequence analysis of Sendai virus F protein has shown the cleavage site to be between positions 116 and 117 (Arg-Phe; Blumberg *et al* ,1985) which is readily cleaved *in vitro* by trypsin. However sequence analysis of various Sendai virus mutants and their revertants has determined that amino acid changes at or close to the cleavage site are responsible for the variation in susceptibility of the F_0 polypeptide to different proteases. Interestingly, a single amino acid change of Arg (116) to Ile not only conferred trypsin resistance but also susceptibility to chymotrypsin. However, a single amino acid change of Phe to Ser at position 117 does not inhibit trypsin digestion but permits digestion with chicken plasmin. Other amino acid changes near the cleavage site are responsible for cleavage by either chymotrypsin, elastase, plasmin and factor X (Itoh *et al*, 1987, Hsu *et al*, 1987, Itoh and Homma, 1988, Ogasawara *et al*, 1992). Thus it was of interest to determine whether the pathogenicity of strain J3666 of PVM could be due to amino acid changes located within its F protein when compared to that of strain 15.

3.2. RESULTS :

3.2.1. Amplification of full-length F-gene of PVM strain J3666 :

First strand cDNA was synthesised from mRNA isolated from BS-C-1 cells infected with PVM strain J3666 as described in the methods section. Amplification of full-length F gene of PVM strain J3666 was carried out by PCR using

oligonucleotides F5 (Table 3.1) and oli-Tail (Table 3.1). Oligo F5 was designed to hybridise to the 5' gene start at positions 1 to 17 and oli-Tail was designed on the basis of the PVM strain 15 sequences to hybridise to the poly A tail. Amplification was achieved using Taq DNA polymerase. Bands of 1450 and 1650 kilo bases (kb) were visualised following agarose gel electrophoresis (Fig. 3.1). The 1650 kb band corresponds to the full-length F gene i.e. amplification following hybridisation of oligo F5 to the 5' gene end and oli-Tail hybridising to the poly A tail. The 1450 kb band results from the 'incorrect' binding of oil-Tail to a stretch of A residues approximately 200 bases upstream of the poly A tail (Data not shown).

Since oligo F5 was designed using the F gene sequence of PVM strain 15, a consequence of this was that the 5' end of the amplified F gene of PVM strain J3666 was identical to that of PVM strain 15. Consequently oligonucleotides G1 and F2 (Table 3.1) were used to amplify the true 5' sequence of the F gene of PVM strain J3666 from polycistronic cDNA (Fig 3.1). Oligo G1 was designed to hybridise to the nucleotide positions 1223 to 1245 at the 3' end of the G gene of PVM strain 15 whereas oligo F2 hybridised to nucleotides 1102 to 1121 on the complementary (-) strand of the F gene of PVM strain 15.

3.2.2. Cloning and sequencing of the F gene PCR products :

The PCR product generated by amplification with oligo's F5 and oli-Tail, was restriction endonuclease digested with *EcoRI* and *HindIII*. The larger of the two bands (Fig. 3.1) corresponding to the whole gene was purified from a low melting point agarose gel and cloned into *EcoRI* and *HindIII* digested cloning vector Bluescribe (pBS). Two individual clones were subcloned into either M13 mp18 or M13 mp19 using a variety of oligonucleotides. Fig. 3.2 shows the strategy adopted for sequencing of PVM strain J3666 F gene. The nucleotide sequence was determined using a range of F gene (PVM strain 15) specific primers shown in Table 3.1. The PCR

Oligo	Sequence	Position	Strand	Usage
G1	GCCTTACTGGTGTCCTAT (G-gene)	1223-1245	P	PCR
Oli-Tail	GGCCCGGGAAGC$(T)_{15}$	Poly A tail*	C	PCR
F2	GTCAGTACCGCTGTACTCAC	1102-1121	C	SP
F4	GCCGGCCTGTTCTAAGGGCACTC	132-149**	C	PCR, SP
F5	GTATACCCGGGAATTCAGGACAA ATATGATTCC	1-17***	P	PCR
F6.5	CATCACGAGAATGCAACTCA	1031-1050	P	SP
F7	TATTGAGTCATGCAAGAGCA	192-211	P	SP
F8	AGGCTGTTGTTAGCCTAACC	431-450	P	SP
F9	CATTGTAGGCGGATGGCTG	681-700	P	SP
F10	GCTAGAGCTGATAATGGCTG	901-920	P	SP
F11	ACTGTTTACTATCTTAGCAA	1261-1280	P	SP
F12	ATATACTAACAACTCTGCTC	1481-1500	P	SP
F13	GTCCAATAACTGATCAGAGG	1440-1421	C	SP
F14	ACAGTCTATAGAACTTCTTA	851-870	C	SP
F15	TGTGTGTTAATCCTGCATTA	621-640	C	SP
F16	GGTTAGGCTAACAACAGCCT	431-450	C	SP
F17	TCCACTGCACTACTATAGATT	240-260	C	SP

Table 3.1 : PVM F gene specific primers and oligonucleotides.

showing nucleotide sequence and position of oligonucleotides and primers used in either PCR or in sequencing reactions. C denotes complementary (-) strand whereas P represents (+) strand sequence; SP - sequencing primer and PCR - oligonucleotide used in the polymerase chain reaction. * : Oli-Tail contains restriction sites for *SmaI* and *HindIII*. ** : Bases underlined are not found within F gene sequence. *** : Oligo F5 contains restriction sites for *KpnI*, *SmaI* and *EcoRI*. Underlined bases are not located within the F gene sequence.

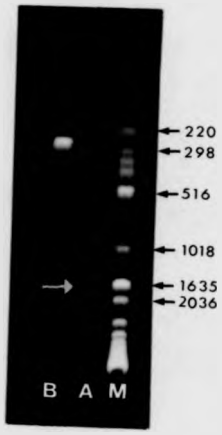


Fig. 3.1 : PVM strain J3666 F gene PCR product.

Low melting point agarose gel stained with ethidium bromide showing a) F gene of PVM strain J3666 amplified by PCR using oligonucleotides F5 and Oli-Tail and b) the PCR product representing the 5' end of PVM strain J3666 F gene following amplification with oligonucleotides F4 and G1. M represents the DNA marker track. The band indicated (arrowed) shows the PCR product purified following amplification with oligonucleotides F5 and Oli-Tail.

product generated by amplification with oligos G1 and F4 was end-filled using T4 DNA polymerase and cloned into M13 mp18. The 5' sequence was determined from six clones.

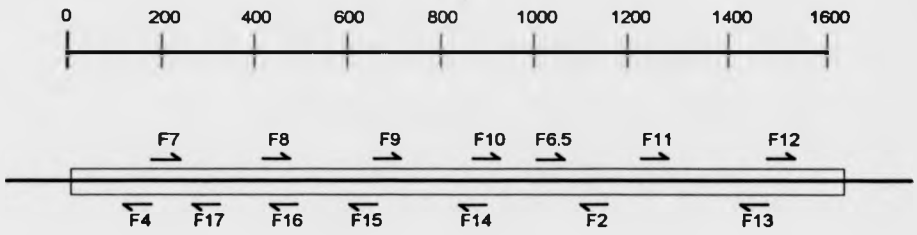


Fig. 3.2 Sequencing strategy for PVM strain J3666 F gene.

Sequencing strategy for PVM strain J3666 F gene with arrows indicating location of oligonucleotides used in sequencing. Sequencing oligonucleotides shown represent their positions in relation to the F gene cDNA. The gene ends were sequenced using the -40 (universal) primer.

3.2.3. Sequence analysis of the F gene of PVM strain J3666 :

Figure 3.3 shows the nucleotide and amino acid changes found in the F gene of PVM strain J3666 compared to the published sequence of Chambers *et al* (1992) for PVM strain 15. As is the case with the F gene of PVM strain 15, the F gene of PVM strain J3666 encodes a large open reading frame of 537 amino acids with a nine base non-coding region at the 5' end and a non-coding region of 37 bases at the 3' end. Only 6 nucleotide changes were found in the whole of the F gene when compared to the F gene sequence of PVM strain 15. All were found within the coding region of the F gene. Table 3.2 summarises the nucleotide and amino acid changes found in the F genes of the two strain of PVM. The substitution of a T residue for an A residue at position 92 results in an amino acid change of a Tyr to Phe whereas the substitution of

```
                    MetIleProGlyArgIlePheLeuValLeuLeuValIlePh
Strain 15/J3666     AGGACAAAUAUGAUUCCUGGCAGGAUCUUUCUAGUCCUUCUGGUGAUCUU  50

                    eAsnThrLysProIleHisProAsnThrLeuThrGluLysTyrTyrGluS
Strain 15           CAACACCAAGCCAAUUCACCCAAAUACAUUAACAGAAAAAUACUAUGAGU 100
Strain J3666                                                  U
                                                             Phe

                    erThrCysSerValGluThrAlaGlyTyrLysSerAlaLeuArgThrGly
Strain 15/J3666     CCACAUGUAGUGUUGAGACUGCAGGUUAUAAGAGUGCCCUUAGAACAGGU 150

                    TrpHisMetThrValMetSerIleLysLeuSerGlnIleAsnIleGluSe
Strain 15/J3666     UGGCAUAUGACAGUUAUGUCAAUUAAGUUGUCUCAAAUAAAUAUUGAGUC 200

                    rCysLysSerSerAsnSerLeuLeuAlaHisGluLeuAlaIleTyrSerS
Strain 15/J3666     AUGCAAGAGCAGCAACUCGUUAUUGGCUCAUGAGCUUGCAAUCUAUAGUA 250

                    erAlaValAspGluLeuArgThrLeuSerSerAsnAlaLeuLysSerLys
Strain 15/J3666     GUGCAGUGGAUGAAUUGAGAACGUUAUCAUCCAAUGCCUUGAAGUCCAAA 300

                    ArgLysLysArgPheLeuGlyLeuIleLeuGlyLeuGlyAlaAlaValTh
Strain 15/J3666     AGGAAGAAGAGGUUCCUCGGUUUGAUUCUUGGUCUCGGAGCUGCAGUCAC 350

                    rAlaGlyValAlaLeuAlaLysThrValGlnLeuGluSerGluIleAlaL
Strain 15/J3666     UGCCGGGGUGGCUUUAGCCAAGACAGUGCAACUUGAAAGUGAGAUUGCAU 400

                    euIleArgAspAlaValArgAsnThrAsnGluAlaValValSerLeuThr
Strain 15/J3666     UGAUUAGAGAUGCAGUGAGAAAUACAAAUGAGGCUGUUGUUAGCCUAACC 450

                    AsnGlyMetSerValLeuAlaLysValValAspAspLeuLysAsnPheIl
Strain 15/J3666     AACGGCAUGUCAGUGUUGGCUAAAGUGGUGGAUGAUUUGAAAAACUUCAU 500

                    eSerLysGluLeuLeuProLysIleAsnArgValSerCysAspValHisA
Strain 15/J3666     AUCUAAAGAAUUACUCCCAAAAAUAAACCGAGUCUCUUGUGAUGUGCACG 550

                    spIleThrAlaValIleArgPheGlnGlnLeuAsnLysArgLeuLeuGlu
Strain 15/J3666     ACAUCACUGCCGUCAUUAGAUUCCAACAGCUCAACAAAAGACUUUUGGAA 600

                    ValSerArgGluPheSerSerAsnAlaGlyLeuThrHisThrValSerSe
Strain 15/J3666     GUAUCUCGUGAAUUUUCAUCUAAUGCAGGAUUAACACACACUGUUUCAUC 650

                    rPheMetLeuThrAspArgGluLeuThrSerIleValGlyGlyMetAlaV
Strain 15           UUUUAUGUUAACAGACCGCGAACUCACCUCCAUUGUAGGCGGCAUGGCUG 700
Strain J3666                  G
                             Arg

                    alSerAlaGlyGlnLysGluIleMetLeuSerSerLysAlaIleMetArg
Strain 15           UUUCAGCAGGCCAAAAAGAGAUAAUGCUAUCUAGCAAAGCUAUAAUGAGA 750
Strain J3666                                    G
                                               Arg
```

```
                ArgAsnGlyLeuAlaIleLeuSerSerValAsnAlaAspThrLeuValTy
Strain 15/J3666 AGAAAUGGGUUAGCAAUAUUAAGUUCAGUCAACGCUGACACACUGGUUUA 800

                rValIleGlnLeuProLeuPheGlyValMetAspThrAspCysTrpValI
Strain 15/J3666 UGUAAUACAACUCCCAUUAUUUGGUGUUAUGGACACAGAUUGUUGGGUAA 850

                leArgSerSerIleAspCysHisAsnIleAlaAspLysTyrAlaCysLeu
Strain 15/J3666 UAAGAAGUUCUAUAGACUGUCAUAACAUAGCAGACAAGUAUGCUUGUUUG 900

                AlaArgAlaAspAsnGlyTrpTyrCysHisAsnAlaGlySerLeuSerTy
Strain 15/J3666 GCUAGAGCUGAUAAUGGCUGGUAUUGUCACAAUGCUGGCUCAUUAUCAUA 950

                rPheProSerProThrAspCysGluIleHisAsnGlyTyrAlaPheCysA
Strain 15       CUUCCCGUCGCCAACGGAUUGUGAGAUCCACAAUGGGUAUGCUUUCUGUG 1000
Strain J3666                  U                               U
                             Thr                             Val

                spThrLeuLysSerLeuThrValProValThrSerArgGluCysAsnSer
Strain 15/J3666 ACACUCUAAAAAGUCUAACUGUACCUGUAACAUCACGAGAAUGCAACUCA 1050

                AsnMetTyrThrThrAsnTyrAspCysLysIleSerThrSerLysThrTy
Strain 15/J3666 AACAUGUAUACCACUAACUACGAUUGUAAGAUUUCCACAAGUAAAACUUA 1100

                rValSerThrAlaValLeuThrThrMetGlyCysLeuValSerCysTyrG
Strain 15/J3666 UGUGAGUACAGCGGUACUGACUACAAUGGGUUGCUUGGUAUCUUGUUAUG 1150

                lyHisAsnSerCysThrValIleAsnAsnAspLysGlyIleIleArgThr
Strain 15/J3666 GUCAUAACAGUUGCACAGUCAUCAAUAAUGACAAAGGUAUAAUAAGGACU 1200

                LeuProAspGlyCysHisTyrIleSerAsnLysGlyValAspArgValGl
Strain 15/J3666 CUGCCAGAUGGUUGCCACUACAUCUCCAACAAAGGUGUGGACAGGGUUCA 1250

                nValGlyAsnThrValTyrTyrLeuSerLysGluValGlyLysSerIleV
Strain 15/J3666 AGUAGGUAACACUGUUUACUAUCUUAGCAAAGAAGUUGGCAAGUCAAUUG 1300

                alValArgGlyGluProLeuValLeuLysTyrAspProLeuSerPhePro
Strain 15/J3666 UUGUCAGAGGGGAACCAUUGGUCUUGAAAUAUGACCCUUUGAGUUUCCCU 1350

                AspAspLysPheAspValAlaIleArgAspValGluHisSerIleAsnGl
Strain 15/J3666 GACGAUAAAUUUGAUGUUGCUAUAAGAGAUGUGGAGCAUAGCAUCAAUCA 1400

                nThrArgThrPhePheLysAlaSerAspGlnLeuLeuAspLeuSerGluA
Strain 15       GACACGCACAUUCUUCAAGGCCUCUGAUCAGUUAUUGGACUUAAGUGAAA 1450
Strain J3666                  G
                             Leu

                snArgGluAsnLysAsnLeuAsnLysSerTyrIleLeuThrThrLeuLeu
Strain 15/J3666 ACAGAGAGAAUAAAAAUUUAAACAAGUCAUAUAUACUAACAACUCUGCUC 1500
```

	PheValValMetLeuIleIleIleMetAlaValIleGlyPheIleLeuTy	
Strain 15/J3666	UUCGUUGUAAUGCUUAUUAUAAUAAUGGCUGUCAUAGGGUUCAUUCUGUA	1550
	rLysValLeuLysMetIleArgAspAsnLysLeuLysSerLysSerThrP	
Strain 15/J3666	UAAGGUAUUGAAAAUGAUCAGAGACAACAAGUUGAAAUCCAAAAGUACAC	1600
	roGlyLeuThrValLeuSer	
Strain 15/J3666	CUGGCCUCACAGUUUUAUCAUGACAAUUGUACCAAACCAUAAUUGAGUUA	1650
Strain 15/J3666	GUUAAUU	1657

Fig. 3.3 : PVM strains 15 and J3666 F gene nucleotide and amino acid sequences.

Nucleotide and amino acid sequence of the F gene of PVM strains 15 and J3666. Nucleotide sequence shown is mRNA sense and represents the sequence of PVM strain 15. Nucleotide bases indicated below the sequence for PVM strain 15 F gene represent nucleotide changes found in the F gene of PVM strain J3666. The amino acid sequence shown below the nucleotide changes identified in the G gene of PVM strain J3666 represent changes in amino acid sequence of PVM strain J3666 F gene compared to that of PVM strain 15. A summary of the nucleotide and amino acid changes is shown in Table 3.2.

Nucleotide Changes			Amino acid changes		
Position	**From**	**To**	**Position**	**From**	**To**
92	TAC	TTC	28	Tyr	Phe
669	CGC	CGG	220	Arg	Arg
737	AAA	AGA	243	Lys	Arg
966	ACG	ACT	319	Thr	Thr
992	GCT	GTT	328	Ala	Val
1416	TTC	TTG	468	Phe	Leu

Table 3.2 : PVM F gene nucleotide and amino acid sequence changes.

Summary of nucleotide and amino acid changes found in the F gene of PVM strain J3666 compared to the F gene of PVM strain 15. Nucleotide codons are represented as mRNA sense cDNA and base changes are shown in an enlarged and underlined font.

G residue for C residue at position 669 and T for a G at amino acid 966 conserves the amino acids. G residue substitutions for an A and G at positions 737 and 1416 result in the amino acid changes of Lys to Arg and Phe to Leu respectively. Substitution of a T residue for C at position 992 results in an amino acid change of Ala to Val.

3.3. DISCUSSION :

Sequence analysis of the fusion (F) gene of a pathogenic strain of PVM (strain J3666) showed only a few nucleotide and amino acid changes compared to the sequence of the fusion gene of a non-pathogenic strain. In total only four amino acid

changes were seen, all of which were conservative and unlikely to have a significant effect on the major structural regions identified previously (Chambers *et al*, 1992).

Correlating the amino acid changes found in the F protein of PVM strain J3666 with the regions identified by Chambers *et al* (1992) reveals that the amino acid change of Tyr (28) to Ser occurs within the F_2h region whereas the Lys to Arg change at position 243 is located within the variable domain of the protein (Fig. 1.3). The Ala (328) to Val change is located within the Cysteine-rich domain and the Phe to Leu change at position 468 occurs within the heptad repeat region located at the C-terminus of the F_1 subunit. The importance of these changes can be tentatively correlated with the antigenic sites identified within the HRSV F protein.

Two approaches to map the antigenic domains of the F protein of HRSV by monoclonal antibodies have been undertaken by researchers. Firstly, the use of synthetic peptides to locate the antigenic domains of the F protein of HRSV have proved somewhat successful. Trudel *et al* (1987) were the first to use such an approach in which synthetic peptides were used to identify the binding site of a monoclonal antibody, 7C2. The binding site was mapped to a stretch of amino acid residues running from position 221 (Ile) to 232 (Glu). This region is located within the heptad region located at the N-terminus of the F_1 subunit identified by Chambers *et al* (1992). However MAb 7C2 has also been shown by dot-blot assays to react with peptides spanning positions 221 to 232 and 275 to 288 (Trudel *et al*, 1991). So the precise epitope is not yet clear. Using similar approaches Bourgeois *et al* (1991) identified two regions involved in antibody binding ; amino acids 220 to 225 and 259 to 278. The first region is within that identified by Trudel *et al* (1987) whereas the second is located within the variable domain (Chambers *et al*, 1992). Scopes *et al* (1990) have also mapped a polyclonal antibody binding to a region of amino acids at position 483 to 488. By replacing certain amino acid within the synthetic peptides they determined that Pro (484) and Glu (487) were crucial in antibody binding. This region

overlaps the C-terminus of the Cysteine-rich domain and the N-terminus of the heptad repeat region (located at the C-terminus) of the F_1 subunit as identified by Chambers *et al* (1992). Martin-Gallardo *et al* (1991) have mapped a virus-neutralising epitope to amino acids 289 to 298. This region maps to the variable domain of the fusion protein. Thus peptide mapping of the fusion protein has identified two regions for antibody recognition - 200 to 298 and 483 to 488, with the first containing multiple antibody recognition sites.

The second approach used to map antibody binding involves raising escapes mutants against neutralising antibodies and then pinpointing the nucleotide and amino acid changes. Lopez *et al* (1990) have adopted this strategy to identify epitopes involved in antibody binding in HRSV. The antigenic binding site of the neutralising MAb, 47F was initially mapped in a Western blot against purified F protein digested with trypsin. The MAb was found to react with 20.5 and 19 kDa fragments. The amino terminus of these fragments were found to be identical and matched the first eight residues of the F_1 subunit. Five escape mutants were raised against MAb 47F and nucleotide sequence analysis revealed two of these mutants to have changed amino acid 262 (Asp) to a Tyr residue whereas the other three mutants had amino acid 268 (Asp) changed to Ile. These results suggest that the 3-D conformation of the F_1 N terminus is important in eliciting antibody response. Arbiza *et al* (1992) adopted a combination of the above two strategies in that firstly neutralisation escape mutants of HRSV were raised against a selection of MAb's - 11, 47F, AK13A2, 7C2, B4, 19 and 20 (note that MAb 7C2 was initially used by Trudel *et al* (1987) in mapping linear epitopes using synthetic peptides). The nucleotide and subsequent amino acid changes in the F gene were pinpointed and the reactivity of synthetic peptides with the changed amino acid(s), to MAb's binding to the same region tested. Escape mutants raised against one MAb (e.g. AK13A2) eliminated reactivity with other MAb's (e.g. 7C2 and B4). Similarly, synthetic peptides made against one MAb showed little or no reactivity

with other MAb's that bound to the same region. Using this approach Arbiza *et al* (1992) identified two regions involved in antibody binding spanning amino acids 262 to 272, the variable region (Chambers *et al*, 1992) and amino acids 422 to 438 (near that identified by Scopes *et al*, 1990) located within the cysteine-rich region identified by Chambers *et al* (1992). However MAb 7C2 failed to react with a synthetic peptides spanning the regions identified by Trudel *et al* (1991). These results suggest a single amino acid changes within the F protein enables the virus to escape MAb neutralisation and that although linear peptides can be used to map antigenic sites, neutralisation escape mutants should be preferred to identify conformational epitopes eliciting antibody response. Yusoff *et al* (1989) obtained similar results in antibody-neutralising mutants in NDV in which the amino acid changes fell into three regions. The first being the "loop" between the F_2 beta-turn domain and F_2h region, the second was the heptad repeat region at the N terminus of the F_1 subunit and thirdly the variable domain of the F_1 subunit (Chambers *et al*, 1992).

Taking the above reports into consideration it is possible that the amino acid changes located within the variable, cysteine-rich and C-terminus heptad repeat regions of the F protein of PVM could reflect antibody driven changes in sequence. However it seems unlikely that these changes account directly for the difference in pathogenicity between PVM strain 15 and strain J3666.

Using a recombinant vaccinia virus Ebata *et al* (1991) expressed the fusion and haemagglutinin-neuraminidase (HN) of human parainfluenza virus type 3 (HPIV3) either independently or together. They were able to show independent expression and correct maturation of both glycoproteins but could only induce fusion, characterised by syncytium formation and lysis of human red blood cells, when cells were co-infected with both the fusion and HN recombinant vaccinia virus. Hu *et al* (1992) have also demonstrated the requirement for the expression of both the F and HN glycoproteins in human parainfluenza virus type 2 (HPIV2). In this study they

infected cells with recombinant vaccinia virus containing (and expressing) T7 RNA polymerase. Infected cells were then transfected with recombinant plasmids containing the F and HN genes. No fusion activity was observed when HPIV2 F gene and HPIV3 HN gene or HPIV3 F gene and HPIV2 HN gene were co-expressed. Additionally by mixing cells independently expressing the F or HN protein the fusion function was regained (Hu *et al*, 1992). These results indicate that for human parainfluenza viruses the F and HN proteins of the same virus type are required for fusion. Morrison *et al* (1991) used a retroviral expression vector to express the fusion protein of a virulent strain of NDV. No cell to cell fusion was observed in cells expressing the F protein even though the F_0 polypeptide was cleaved into its F_1 and F_2 subunits. Subsequent transfection of the HN gene into these F protein expressing cells elicited cell to cell fusion. Unlike the findings of Hu *et al* (1992) fusion was only observed when both the HN and F proteins were expressed in the same cell. Sergel *et al* (1993) used Cos-7 cells and a SV40 based vector (pSVL) to demonstrate similar results for NDV HN and F proteins. Wild *et al* (1991) and Tanabayashi *et al* (1992) demonstrated a requirement for co-expression of the mumps H and F proteins to elicit cell to cell fusion in vaccinia recombinants. These findings contradict the work of Alkhatib *et al* (1990), who demonstrated cell to cell fusion by expression of the mumps virus fusion protein in adenovirus recombinants. Work on Sendai virus (Bagai *et al*, 1993) has shown that although the fusion protein alone is responsible for fusion, the presence of the HN glycoprotein enhances this activity.

These findings are somewhat complicated by the results of Peeples and Bratt (1984) who found that a mutation in the matrix protein of NDV resulted in the decreased fusion activity. However this work did not test the effect of the presence or absence of the NDV HN glycoprotein upon fusion activity. Thus, the above observations suggest a crucial role for the attachment glycoprotein in not only being responsible for attachment of the virus to the cell but also, in some paramyxoviruses,

to be involved in the fusion process as well. The changes observed in the PVM strain J3666 F protein may affect interaction of the F and G proteins. If the *Pneumovirus* F protein requires the presence of the G protein for activity any alterations in F may indirectly affect fusion activity. Thus, it remains possible that one or more of the observed changes in PVM strain J3666 F protein may be important for enhanced pathogenicity. Thus it would be interesting to determine at the molecular level whether the attachment (G) glycoprotein of PVM plays a role in viral pathogenicity and fusion.

CHAPTER 4

NUCLEOTIDE SEQUENCE OF PVM G GENES.

4.1. INTRODUCTION :

The G glycoprotein of HRSV has been shown to be responsible for virus attachment to the cell surface (Levine *et al*, 1987). In the preceding chapter the fusion protein of the pathogenic strain of PVM, J3666, was shown to possess only minor changes when compared to that of a non-pathogenic strain that has been passaged entirely in tissue culture. This lack of sequence variation suggests that the G protein may be important in determining pathogenicity. Although the fusion proteins of HRSV viruses are well conserved and have been demonstrated by the use of recombinant vaccinia virus and purified F protein to elicit production of neutralising antibodies that protect immunised mice against HRSV challenge, a dilemma exists as to why HRSV is able to re-infect individuals. Work by Anderson *et al* (1988) suggests a mixture of anti- F and anti- G neutralising MAb's to be the key to this question. It is likely to be the G protein that is important in the ability of HRSV to re-infect individuals who have previously been infected and have good anti-HRSV antibody response. Sequence analysis of several HRSV G genes confirms the divergent nature of this protein (Anderson *et al*, 1985, Akerlind *et al*, 1988, Cane *et al*, 1991, Garcia-Barreno *et al*, 1989, Sullender *et al*, 1991). Furthermore the generation of HRSV escape mutants raised against neutralising anti-G MAb's suggest several types of modifications may be employed by the virus in the evolution of the G protein in response to immunological pressures. Escape mutants have been found to be generated by single amino acid substitutions, frameshifting due to insertion/deletion of nucleotides or the presence of premature stop codons (Garcia-Barreno *et al*, 1990, Rueda *et al*, 1991). Cane *et al* (1991) identified conserved and hypervariable regions within the G gene of HRSV natural isolates. The cytoplasmic and transmembrane regions were found to be conserved between these isolates whereas the extracellular domain consisted of alternating conserved and variable regions. The C terminus of the G protein was found to be variable with a hydrophilic hypervariable region located

between amino acids 101 and 133 with amino acid differences ranging from 2-48%. These results support the findings of Garcia-Barreno *et al* (1990) who determined frameshift mutations in HRSV mutants that escape neutralisation against an anti-G MAb to be localised to the C terminus of the HRSV G protein. Thus, variation within the attachment protein of HRSV appears to be driven by immunological pressures. For these reasons, and the apparent lack of significant sequence differences between the F genes of the pathogenic and non-pathogenic strains of PVM, sequence determination of the G proteins may be useful in shedding light on the molecular basis for the differing pathogenicities observed for these two PVM strains.

4.2. RESULTS :

4.2.1. Sequence analysis of the G gene of PVM strain J3666 :

4.2.1.1. CLONING AND SEQUENCING OF THE G GENE :

Amplified PVM strain J3666 G gene was kindly provided by Dr. J. Johnson. The G gene was amplified from cDNA synthesised from mRNA from PVM strain J3666 infected cells by PCR using oligonucleotide G4 (CCCAAGCTTGGCCAGGATAAGTACTATCC, designed to hybridise to the 5' mRNA end of the G gene at positions 1 to 16 (of the nucleotide sequence shown in Fig. 4.4) and containing restriction endonuclease sites for *MscI*, *HindIII*) and oligonucleotide oli-Tail (GGCCCGGGAAGC$(T)_{15}$, designed to hybridise to the poly A tail of the mRNA and containing restriction endonuclease sites for *SmaI*). The PCR product was digested with restriction endonucleases *SmaI* and *MscI*, purified from a low melting point agarose gel (Fig. 4.1) and cloned into the *SmaI* restriction site of the plasmid pBS (Bluescribe; Stratagene, La Jolla, CA, USA)

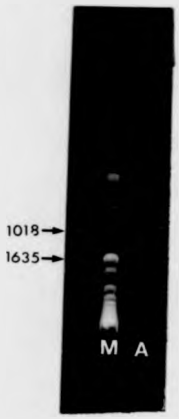


Fig. 4.1: PVM strain J3666 G gene PCR product.

Agarose gel stained with ethidium bromide showing the G gene of PVM strain J3666 (A) following amplification by PCR in the presence of Vent™ DNA polymerase and oligonucleotides G4 and oli-Tail. The G gene was kindly provided by Dr. J. Johnson. M represents the DNA marker lane.

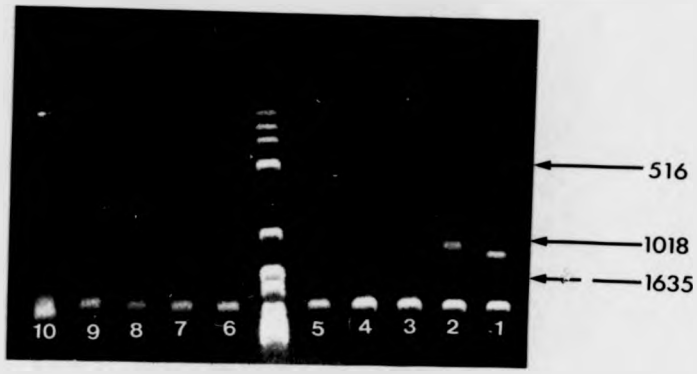
The 5' ends of two independent clones containing the G gene of strain J3666 that were in opposite orientations to each other were linearised with either *PstI* or *EcoRI* and digested with exonuclease *Bal31*. At regular intervals aliquots were taken out, and the reaction stopped by the addition of EDTA to a final concentration of 25mM. T4 DNA polymerase was used to blunt-end the *Bal31* digested DNA as described in chapter 2. The truncated fragments were released by digestion with *EcoRI* or *PstI* and purified from a low melting point agarose (Fig. 4.2) prior to cloning into *SmaI* and *EcoRI* or *PstI* digested pBS. A selection of various sizes of these truncated inserts were subcloned into M13 mp18 or mp19 sequencing vectors following

estimation of insert sizes on a 1% agarose gel of a number of *EcoRI* and *PstI* digested DNA prepared by the miniprep protocol. Nucleotide sequence of PVM strain J3666 was determined using the commercially available Sequenase® kit. The deletions within the G gene of PVM strain J3666 were created not just to determine the nucleotide sequence of the gene but also for the potential use of such deletions in the mapping of antigenic sites.

4.2.1.2. CONFIRMATION OF THE 5' END NUCLEOTIDE SEQUENCE :

A consequence of the PCR amplification procedure adopted above for cloning of the PVM strain J3666 G gene was that the 5' end sequence would represent the sequence of oligo G4 used in the PCR reaction. To determine the correct sequence of the 5' end of the G gene, amplification of the 5' end of the G gene was then achieved by PCR with the use of oligonucleotide 1A4 (CCCCGGGCCAGACAGCTTTACCC), designed to hybridise to the 3' end of the PVM 1A gene, and oligo G3 (GTGAGCACACCTGCAATAAGCTTA) designed to hybridise to nucleotides 188 to 211 of the sequence determined above (Fig. 4.3). The PCR product was digested with *SmaI* and *HindIII* (restriction sites located within oligonucleotides Oli-C and G3 respectively) and cloned into M13 mp18. Nucleotide sequence was confirmed using the commercially available Sequenase® kit.

a)



b)

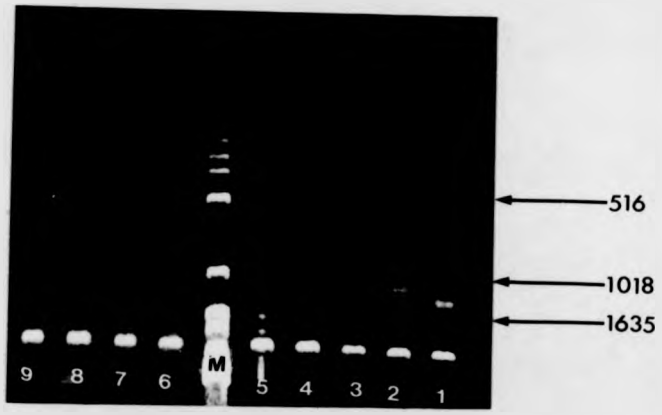


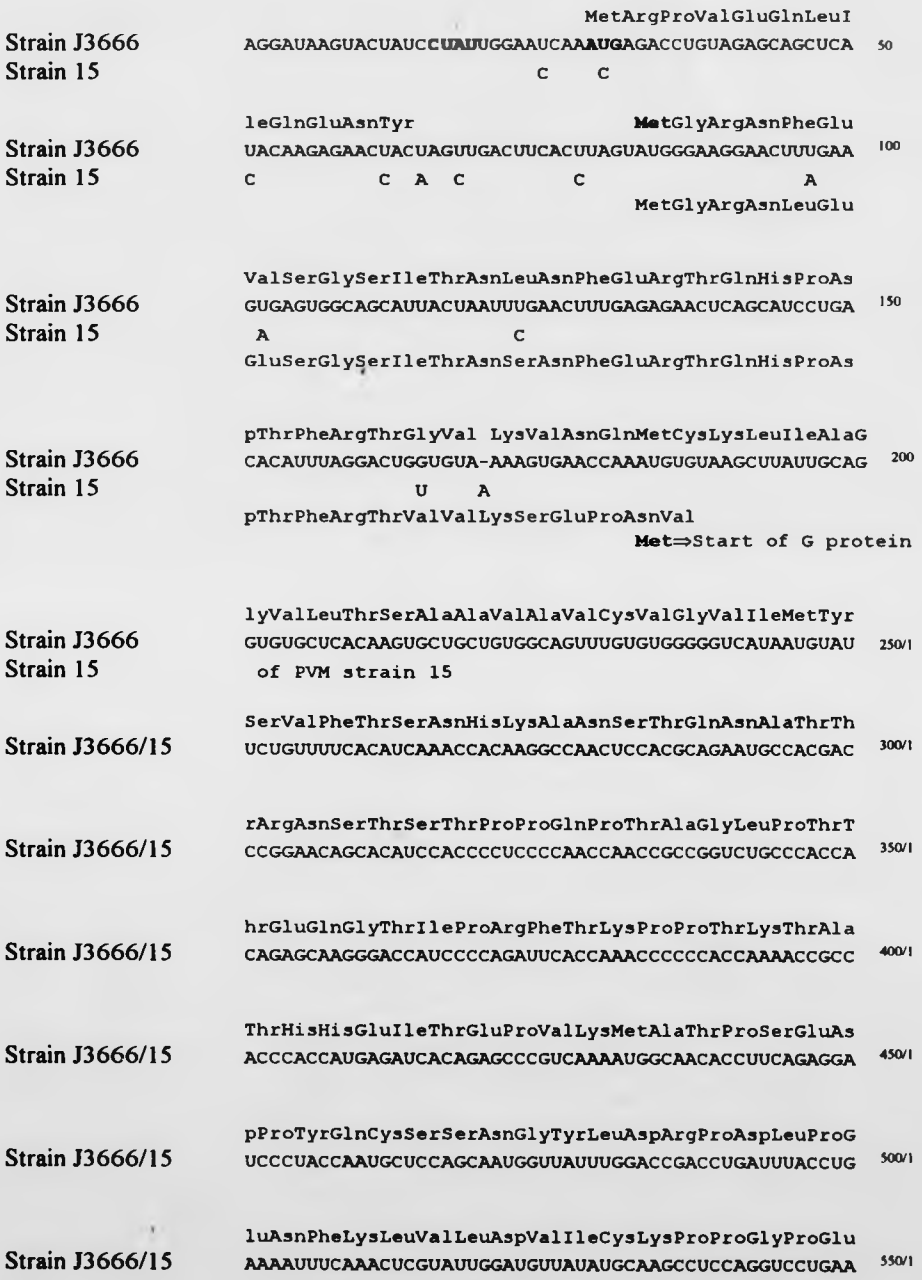
Fig. 4.2 : PVM strain J3666 G gene truncations.

Low melting point agarose gel stained with ethidium bromide showing 3' and 5' end truncated fragments of the G gene of PVM strain J3666 following linearisation with either a) *EcoRI* or b) *PstI*, digestion with exonuclease *Bal31*, end-blunted with T4 DNA polymerase and released with digestion with either *EcoRI* or *PstI*. Numbers 1-10 represent intervals of 5, 15, 30, 45, 60, 75, 90, 105, 120 and 135 mins at which aliquots of the *Bal31* reaction were stopped by the addition of EDTA. Purified fragments were ligated into *Sma1* and *EcoRI* or *PstI* cut pBS.

Fig. 4.3 : PCR product representing PVM strain J3666 5' end.

1% agarose gel stained with ethidium bromide showing the PCR product (A) following amplification with oligonucleotides 1A4 and G3. M represents the DNA marker lane.

Fig 4.4 shows the nucleotide and predicted amino acid sequence of the G gene of PVM strains J3666 and 15 (See later). PVM strain J3666 G gene is 1329 nucleotides long and contains two non-overlapping open reading frames (ORF's), which are diagrammatically shown in Fig. 4.9. The first ORF codes for a small polypeptide of 12 amino acids whereas, as is the case for HRSV G gene, the second ORF codes for the major polypeptide of 396 amino acids with a predicted molecular weight of 43,584. The gene has 6 potential N linked glycosylation sites (indicated in Fig. 4.4) and a high percentage of serine, threonine and proline residues, which are potential sites for the O-linked glycosylation characterised by Ling and Pringle (1989b).


MetArgProValGluGlnLeuI
Strain J3666
AGGAUAAGUACUAUCCUAUUGGAAUCAAAUGAGACCUGUAGAGCAGCUCA 50
Strain 15
C C
leGlnGluAsnTyr MetGlyArgAsnPheGlu
Strain J3666
UACAAGAGAACUACUAGUUGACUUCACUUAGUAUGGGAAGGAACUUUGAA 100
Strain 15
C C A C C A
MetGlyArgAsnLeuGlu
ValSerGlySerIleThrAsnLeuAsnPheGluArgThrGlnHisProAs
Strain J3666
GUGAGUGGCAGCAUUACUAAUUUGAACUUUGAGAGAACUCAGCAUCCUGA 150
Strain 15
A C
GluSerGlySerIleThrAsnSerAsnPheGluArgThrGlnHisProAs
pThrPheArgThrGlyVal LysValAsnGlnMetCysLysLeuIleAlaG
Strain J3666
CACAUUUAGGACUGGUGUA-AAAGUGAACCAAAUGUGUAAGCUUAUUGCAG 200
Strain 15
U A
pThrPheArgThrValValLysSerGluProAsnVal
Met⇒Start of G protein
lyValLeuThrSerAlaAlaValAlaValCysValGlyValIleMetTyr
Strain J3666
GUGUGCUCACAAGUGCUGCUGUGGCAGUUUGUGUGGGGGUCAUAAUGUAU 250/1
Strain 15
of PVM strain 15
SerValPheThrSerAsnHisLysAlaAsnSerThrGlnAsnAlaThrTh
Strain J3666/15
UCUGUUUUCACAUCAAACCACAAGGCCAACUCCACGCAGAAUGCCACGAC 300/1
rArgAsnSerThrSerThrProProGlnProThrAlaGlyLeuProThrT
Strain J3666/15
CCGGAACAGCACAUCCACCCCUCCCCAACCAACCGCCGGUCUGCCCACCA 350/1
hrGluGlnGlyThrIleProArgPheThrLysProProThrLysThrAla
Strain J3666/15
CAGAGCAAGGGACCAUCCCCAGAUUCACCAAACCCCCCACCAAAACCGCC 400/1
ThrHisHisGluIleThrGluProValLysMetAlaThrProSerGluAs
Strain J3666/15
ACCCACCAUGAGAUCACAGAGCCCGUCAAAAUGGCAACACCUUCAGAGGA 450/1
pProTyrGlnCysSerSerAsnGlyTyrLeuAspArgProAspLeuProG
Strain J3666/15
UCCCUACCAAUGCUCCAGCAAUGGUUAUUUGGACCGACCUGAUUUACCUG 500/1
luAsnPheLysLeuValLeuAspValIleCysLysProProGlyProGlu
Strain J3666/15
AAAAUUUCAAACUCGUAUUGGAUGUUAUAUGCAAGCCUCCAGGUCCUGAA 550/1

	HisHisAsnThrSerCysTyrGluLysArgGluIleAsnProGlySerVa	
Strain J3666/15	CAUCACAACACCAGCUGUUAUGAGAAACGUGAAAUCAACCCAGGAAGUGU	600/1
	lCysProAspLeuValThrMetLysAlaAsnMetGlyLeuAsnAsnGlyG	
Strain J3666/15	UUGCCCUGAUCUUGUAACAAUGAAGGCAAACAUGGGCUUAAACAAUGGUG	650/1
	lyGlyGluAspAlaAlaProTyrIleGluValThrThrLeuSerThrTyr	
Strain J3666/15	GUGGGGAGGAUGCUGCACCUUAUAUAGAGGUUACCACCCUUUCUACGUAC	700/1
	SerAsnLysArgAlaMetCysValHisAsnGlyCysAspGlnGlyPheCy	
Strain J3666/15	UCCAACAAAAGGGCAAUGUGUGUCCACAAUGGGUGUGAUCAAGGCUUCUG	750/1
	sPhePheLeuSerGlyLeuSerThrAspGlnGluArgAlaValLeuGluL	
Strain J3666/15	UUUCUUCCUUUCUGGUUUAAGCACUGAUCAGGAGAGAGCUGUGCUAGAGC	800/1
	euGlyGlyGlnGlnAlaIleMetGluLeuHisTyrAspSerTyrTrpLys	
Strain J3666/15	UUGGAGGUCAACAGGCUAUCAUGGAGUUGCAUUAUGAUUCCUACUGGAAA	850/1
	HisTyrTrpSerAsnSerAsnCysValValProArgThrAsnCysAsnLe	
Strain J3666/15	CACUAUUGGAGUAACUCUAAUUGUGUUGUUCCCAGAACAAACUGCAACCU	900/1
	uThrAspGlnThrGluIleLeuPheProArgPheAsnAsnLysAsnGlnS	
Strain J3666/15	GACAGACCAAACUGAGAUUUUGUUUCCUAGGUUUAACAACAAGAAUCAGU	950/1
	erGlnCysThrThrCysAlaAspSerAlaGlyLeuAspAsnLysPheTyr	
Strain J3666	CUCAGUGUACCACCUGUGCAGAUUCAGCUGGCCUAGAUAACAAAUUUUAU	1000/1
Strain 15	G	
	Glu	
	LeuThrCysAspGlyLeuLeuArgThrLeuProLeuValGlyLeuProSe	
Strain J3666/15	CUCACAUGUGAUGGGCUUUUAAGAACCCUCCCUCUAGUUGGACUACCCAG	1050/1
	rLeuSerProGlnAlaTyrLysAlaValProThrGlnThrThrGlyThrT	
Strain J3666/15	CCUAAGUCCUCAGGCUUACAAAGCUGUACCCACACAAACUACAGGCACCA	1100/1
	hrThrAlaProThrSerGluSerArgHisProThrProAlaProArgArg	
Strain J3666	CCACGGCACCAACAUCAGAGUCGAGGCACCCAACCCCUGCACCCAGGAGG	1150/1
Strain 15	A	
	Thr	

	SerLysProLeuSerArgLysLysArgAlaLeuCysGlyValAspSerSe	
Strain J3666/15	UCCAAACCUCUCAGUCGGAAGAAGAGAGCUUUAUGUGGUGUAGACUCAAG	1200/1
	rArgGluProLysProThrMetProTyrTrpCysProMetLeuGlnLeuP	
Strain J3666/15	CAGAGAACCCAAACCAACAAUGCCUUACUGGUGUCCUAUGCUCCAAUUAU	1250/1
	heProArgArgSerAsnSer	
Strain J3666/15	UUCCAAGGAGGUCUAAUUCUUAAGUGACCUAUUCCUGAAUUAACUUCAGA	1300/1
Strain J3666/15	AUAAGUACCAACCUUAUCAGUAGUUAAUG	

Fig. 4.4 : PVM strains J3666 and 15 G gene nucleotide and amino acid sequences.

Nucleotide and amino acid sequence of the G gene of PVM strains 15 and J3666. Nucleotide sequence shown is mRNA sense and represents the sequence of PVM strain J3666. The amino acid sequence of the first and second open reading frames (ORF) is shown above this sequence. Nucleotide bases indicated below the sequence for strain J3666 G gene represent nucleotide changes found in the G gene of PVM strain 15. The amino acid sequence shown below the nucleotide changes identified in the G gene of PVM strain 15 represent changes in amino acid sequence of the two ORF's of PVM strain 15 G gene compared to that of PVM strain J3666. A gap (indicated by a - sign) was introduced at position 171 of the nucleotide sequence of the G gene of strain J3666 to indicate the addition of an A residue in the nucleotide sequence of PVM strain 15.

4.2.2. Sequence analysis of the G gene of PVM strain 15 :

The sequence of the G gene of PVM strain 15 had been determined from two overlapping clones, designated 702 and 321, representing the nucleotide sequence spanning from positions 1 to 950 and 360 to 1330 shown in Fig. 4.4 respectively (P. Chambers. personal communication). Additionally, two other clones 843 and 221 contained the nucleotide sequence from positions 47 to 350 and 1073 to 1330 respectively. Digestion of these clones with various restriction endonucleases generated fragments which were subcloned into either bacteriophage M13 mp18 or mp19. The nucleotide sequence of the G gene was determined using the method of Sanger *et al* (1980) (P. Chambers, personal communication). A full length clone of the G gene was constructed following digestion of both fragments with *HincII*, located within the gene sequence and other restriction sites found within the plasmid vector into which the G gene fragments had been cloned.. Full length clones were selected for by the ability to digest these constructs with *HincII* (P. Chambers. personal communication). However confirmation that the constructs were truly representative of the full length clones was required from nucleotide sequencing. Furthermore, the 5' end of the G gene of PVM strain 15 needed to be verified since the sequence derived from the clone 702 had an inverted 5' sequence resulting from the cloning procedure adopted by Barr *et al* (1991).

4.2.2.1. DETERMINATION OF THE 5' END SEQUENCE :

Barr *et al* (1991) suggested an explanation for the presence of inverted repeat sequences at the 5' ends of the clones representing PVM strain 15 genes. Fig. 4.5 is a diagrammatic representation of the proposed model for generation of 5' inverted sequences. During first strand cDNA synthesis a hairpin loop structure may have formed at the 5' end. Since DNA polymerase I, which possesses 5'→3 exonuclease activity, and not nuclease S1, was used to resolve such structures second

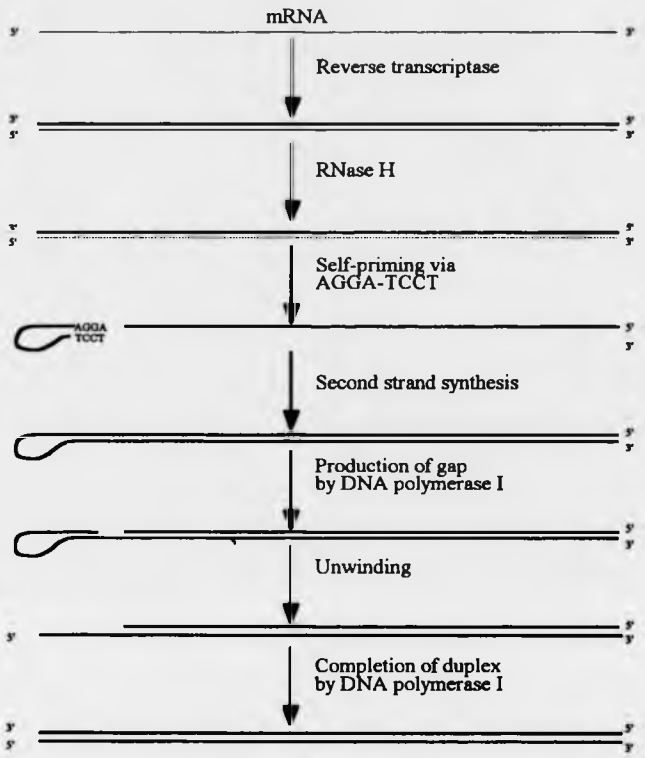


Fig. 4.5 : Model for generation of inverted repeat sequences at PVM strain 15 5' cDNA termini.

Diagrammatic representation of the proposed model for generation of inverted repeat sequences at the 5' end of PVM strain 15 clones (taken from Barr *et al*, 1991). First strand cDNA is synthesised in the presence of AMV reverse transcriptase and the mRNA strand removed following digestion with RNase H. Following second strand cDNA synthesis by DNA polymerase I, a gap is introduced into the first strand at the hairpin loop structure by the 5'→3' exonuclease activity of DNA polymerase I which precedes its 5'→3' polymerase activity. Replication of the loop and stem regions could then occur following unwinding of the molecule resulting in the creation of inverted repeat sequences at the 5' end.

strand cDNA synthesis may have been primed from this hairpin structure.

Following second strand cDNA synthesis, the first strand cDNA synthesised may have been replaced by the action of DNA polymerase 5'→3' exonuclease and 5'→3' polymerase activities. It was argued that during such an event the hairpin loop may represent a physical barrier to further DNA synthesis and that since the DNA polymerase I 5'→3' exonuclease precedes the 5'→3' polymerase activity by a few nucleotides, a gap could have been generated at the end of the duplex. Subsequent unwinding of the nicked loop structure could lead to DNA replication of the loop and stem regions, leading to the creation of inverted repeats at the 5' end. Thus it was necessary to determine the precise 5' sequence of the G gene of PVM strain 15.

A PCR method similar to that used for the 5' end of the N gene by Barr *et al* (1991) was used to amplify the 5' end of the G gene from PVM strain 15 infected cell mRNA in order to determine the precise 5' end sequence of the G gene. Homopolymer G residues were added to the 5' ends of cDNA synthesised from mRNA from PVM strain 15 infected cells in the presence of T4 DNA polynucleotide kinase. The 5' end of PVM strain 15 G gene was amplified by PCR using oligonucleotides oli-C (CCCGAGCTCTGCAGGATCCCCCCCCCCCC) and G3. The PCR product was digested with *BamHI* and *HindIII* (restriction sites located within oligonucleotides oli-C and G3 respectively) and cloned into M13 mp19. Seventeen independent M13 mp19 phage clones containing PCR amplified DNA were selected. Two of these were sequenced individually and the remaining fifteen pooled together and sequenced collectively. The sequence obtained is shown in Fig. 4.6 and was found to be identical to that suggested by Chambers et al (1990b).

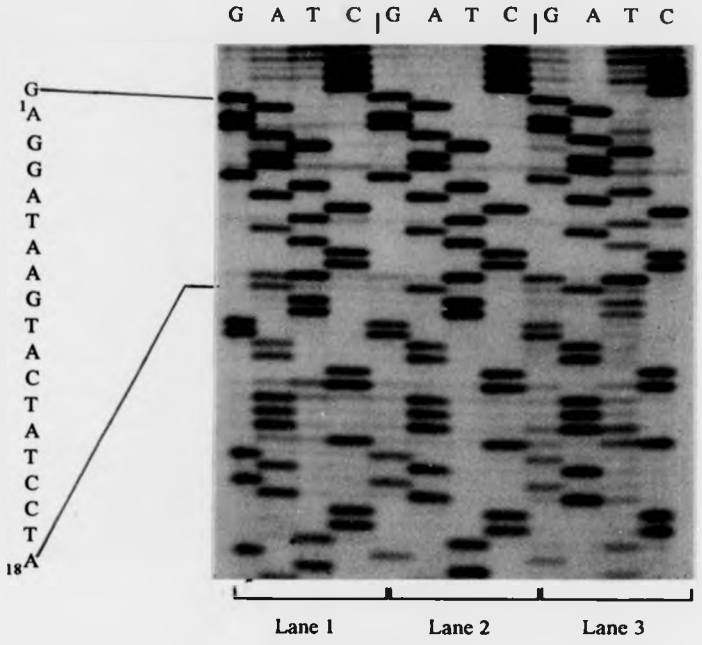


Fig. 4.6 : Nucleotide sequence of 5' end of PVM strain 15 G gene.

Autoradiograph showing the 5' nucleotide sequence of the G gene of PVM strain 15. The sequence shown is in mRNA sense. Lanes 1 and 2 are individual cDNA clones and lane 3 is a pool of 15 independent cDNA clones. The G residue preceding the AGGA sequence probably represents the cap of the mRNA. Numbers shown against individual nucleotides correspond to nucleotide positions shown in Fig. 4.4.

Fig. 4.7 : PVM strain 15 G gene truncations.

Low melting point agarose gel stained with ethidium bromide showing 1) 5' and 2) 3' end truncated fragments of the G(3) clone of PVM strain 15. Truncated fragments were released following end-filling of linearised, *Bal31* digested DNA. Numbers A-H represent intervals of 5, 15, 30, 45, 60, 75, 90 and 105 mins at which aliquots of the *Bal31* reaction were stopped by the addition of EDTA. M represents the DNA marker lane.

4.2.2.2. CONFIRMATION OF FULL LENGTH G GENE CLONES :

CsCl purified DNA from a full length clone, designated G3, was linearised at the 5' end with *EcoRI* and at the 3' end with *PstI*, digested with exonuclease *Bal31*, cloned into the cloning vector pUC13 and subcloned into M13 mp19. Fig. 4.7 shows *Bal31* digested DNA of clone G3 purified from a low melting point agarose gel. Nucleotide sequencing of these truncated fragments confirmed the entire sequence of the G gene for PVM strain 15 (Fig. 4.4).

The nucleotide sequence of the G gene of PVM strain 15 when compared to that of strain J3666 shows several nucleotide changes. As was found for

strain J3666 the G gene of strain 15 also contains two ORF's at the 5' end of the gene. The first ORF potentially encodes a polypeptide of 35 amino acid residues whereas the second ORF codes for the major polypeptide (Fig. 4.9). The second ORF of strain 15 G gene is 1089 nucleotides long coding for a polypeptide of 363 amino acids with a predicted molecular weight of 39,836 (which is close agreement to the molecular weight observed for the unglycosylated G protein of PVM strain 15 determined on SDS-PAGE (Ling and Pringle, 1989b), which is somewhat smaller by 33 amino acids than that of strain J3666. Only two nucleotide changes were found within the coding region for the putative extracellular domain. Nucleotide change of a U residue to A residue resulted in an amino acid change of Phe to Leu at position 97 (Fig. 4.4). A nucleotide change of a C residue to a G residue at position 953/4 (Fig. 4.4) resulted in amino acid substitution of Gln was changed to Glu whereas nucleotide substitution of a U residue to a A residue at position 1121/2 (Fig. 4.4) resulted in an amino acid substitution of Ser to Thr. Nucleotide changes of a U residue to A residue and U residue to C residue at positions 101 and 123 respectively resulted in amino acid changes of Val to Glu and Leu to Ser respectively. The most significant difference observed between the nucleotide sequences of the G genes of the two PVM strains was the addition of an A residue within PVM strain 15 G gene in the stretch of A residues spanning positions 169 to 172 of the G gene of strain J3666. This results in the deletion of 33 amino acid residues (in strain 15) at the N-terminus of the polypeptide synthesised from the major ORF of the G gene strain J3666. The number and position of potential N-link glycosylation sites (Asn-X-Ser/Thr) were conserved in both strains. No significant homology was found between the G genes of PVM strains 15 or J3666 and HRSV virus G gene at both the nucleotide and amino acid level.

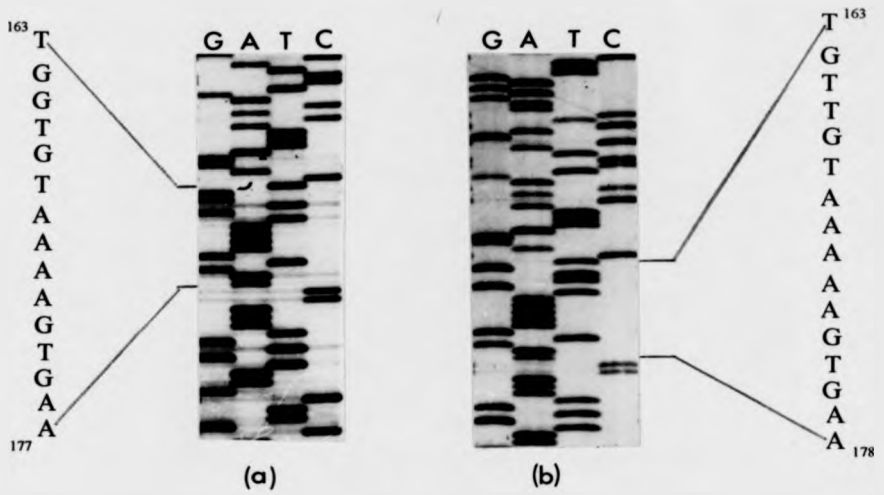


Fig. 4.8 : Nucleotide sequence confirmation of PVM strain J3666 and 15 G gene between positions 163 and 177/8.

Nucleotide sequence showing the presence of a) four and b) five A residues within the G genes of PVM strains J3666 and 15 respectively. The nucleotide sequence shown is located between nucleotide positions 163 and 177/8 of the sequence shown in Fig. 4.4.

To confirm that the addition of the A residue within the G gene of strain 15 between positions 169 and 172 of the sequence for the G gene of strain J3666 was not due to an error during the cloning procedure, a fragment containing the 5' terminal 211 nucleotides of the G gene of PVM strain 15 was amplified using oligonucleotides G3 and G4 from cDNA synthesised from mRNA isolated from PVM strain 15 infected cells (Fig. 4.8). This procedure was deemed crucial since the absence of this additional A residue would result in the addition of 33 amino acids to the N terminus of the polypeptide synthesised from the G gene of PVM strain 15.

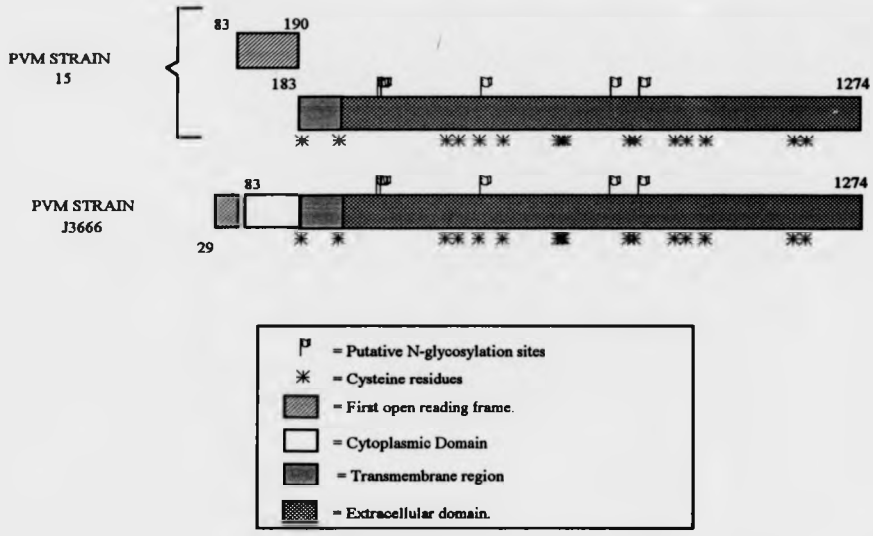


Fig. 4.9 : Open reading frames of PVM strains J3666 and 15 G genes.

Schematic representation of the two open reading frames of the G gene of PVM strains 15 and J3666. Numbers indicate the nucleotides at the start and end of each ORF. Positions of cysteine residues and potential N-linked glycosylation sites are indicated and were conserved in both strains.

Amplification was performed using Taq DNA polymerase as described in chapter 2. Sequence analysis of seventeen independent clones containing the region between nucleotides 75 and 190 was shown to be identical to that obtained above (Fig. 4.8).

A summary of the structure of the G genes of the two strains of PVM is shown diagrammatically in Fig. 4.9. In strain J3666 a 5'-proximal AUG codon is located at positions 29 to 31. This is the beginning of a small ORF capable of encoding a polypeptide of 12 amino acids. In strain 15 a mutation of U to C at position 30 has removed this AUG codon.

4.3. DISCUSSION :

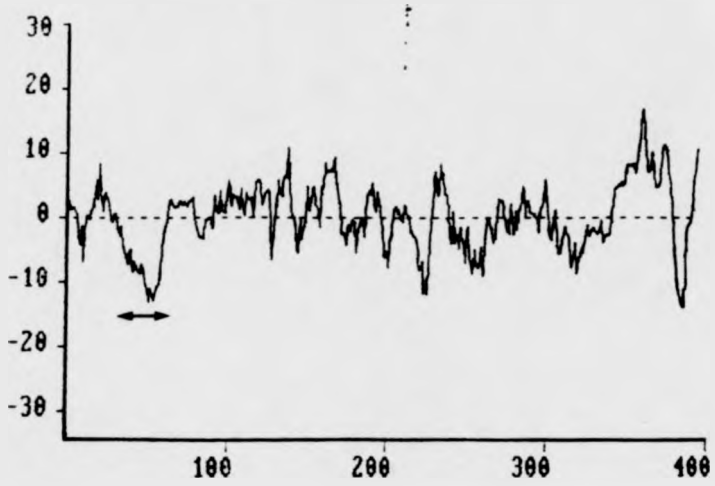
Sequence analysis of the F genes for these two strains, one a highly pathogenic variant and the other a non-pathogenic virus, revealed little nucleotide and amino acid differences (Chapter 3). These changes were considered to be conservative and therefore unlikely to be important in determining the pathogenicity of these strains. Nucleotide sequence of the G genes of these two strains of PVM has been determined. Although HRSV natural isolates have been shown to contain conserved and hypervariable regions within the extracellular domains and conservation within the cytoplasmic and transmembrane regions of the HRSV G proteins (Cane *et al*, 1991), the presence of only two amino acid changes observed within the extracellular domains of the predicted PVM G proteins is probably not a result of changes due to selective pressures but is more likely to be the result of the different passaging protocols used for the two strains since greater sequence divergence was observed within the cytoplasmically located domain of the PVM G proteins. PVM strain J3666 has to our knowledge been entirely passaged in mice whereas PVM strain 15 has been passaged in tissue culture since at least the late 1960's.

The nucleotide and amino acid sequence of the G genes of either of the two strains of PVM show little homology with the G genes of either of the two subgroups of HRSV (Wertz *et al*, 1985; Johnson *et al*, 1987b; Sullender *et al*, 1990). However this is not surprising since the G genes of HRSV differ not only between the two subgroups but also between viruses within the same subgroup (Anderson *et al*, 1985, Akerlind *et al*, 1988, Cane *et al*, 1991, Garcia-Barreno *et al*, 1989, Sullender *et al*, 1991). The predicted polypeptides of 396 and 363 amino acids for PVM strain J3666 and strain 15 are similar in size to that of TRTV (Ling *et al*, 1992) but larger than those of HRSV. Since HRSV lacks demonstrable haemagglutination activity, the concatenation of a haemagglutinin like molecule to the attachment protein of a HRSV like protein, as suggested for other paramyxoviruses cannot be ruled out (Blumberg *et*

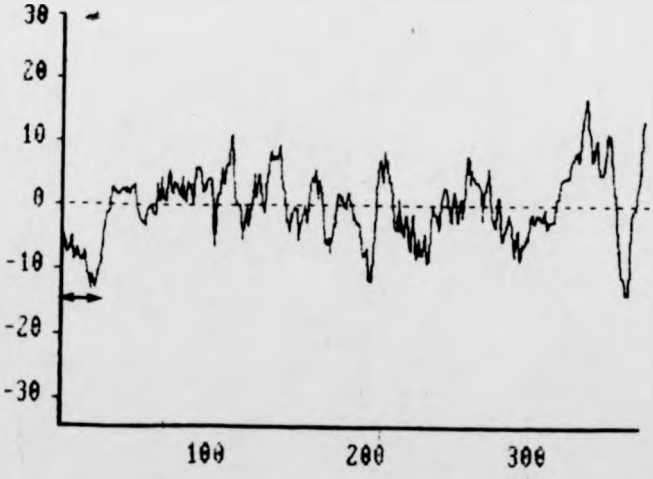
al, 1985, Morrison and Portner, 1991). However, hydropathy analysis reveals some structural features common to the G proteins of pneumoviruses, most notably being the presence of a large hydrophobic domain at the N terminus of the protein (Fig. 4.10). This region has been shown to contain both the signal sequence and membrane anchor region in HRSV (Lichtenstein *et al*, 1991). The amino acid sequence of the major polypeptides synthesised from the G genes of PVM strains 15 and J3666 have approximately equivalent serine (8%) and proline (9%) content but approximately half the threonine (11%) content of that of HRSV (strain A2). However the number of N-linked glycosylation sites found in the G genes of both PVM strains (6 in total) exceed those found in the G gene (4 in total) of HRSV (strain A2; Wertz *et al*, 1985). This is consistent with the presence of high proportion of N- and O-linked glycosylation observed for the G protein of PVM (Ling and Pringle, 1989b).

Another feature of the G genes of both strains of PVM and HRSV (Wertz *et al*, 1985) is the coding of the major polypeptide from the second available AUG. Although the major G polypeptide of strain J3666 is synthesised from the initiation codon, that in strain 15 would code for the small polypeptide of 33 amino acids, another initiation codon is located upstream of this AUG in strain J3666 that would code for another smaller polypeptide of 12 amino acids (Fig. 4.4). Thus, the presence of two initiation codons and the fact that the major polypeptides of the G genes of PVM strains J3666 and 15 are synthesised from the second available AUG, but that the G gene of TRTV is transcribed from the first available open reading frame, suggest a role for the first initiation codons. Although the functional significance of the two open reading frames in PVM and HRSV are not known, one possibility is that the first AUG may act in down regulation of G gene expression since most ribosomes would initiate translation from the 5'-proximal position. Since the gene order of TRTV is different to that of PVM and HRSV in that the G gene is located further from the

A) PVM strain J3666 G protein :



B) PVM strain 15 G protein :



C) HRSV G protein

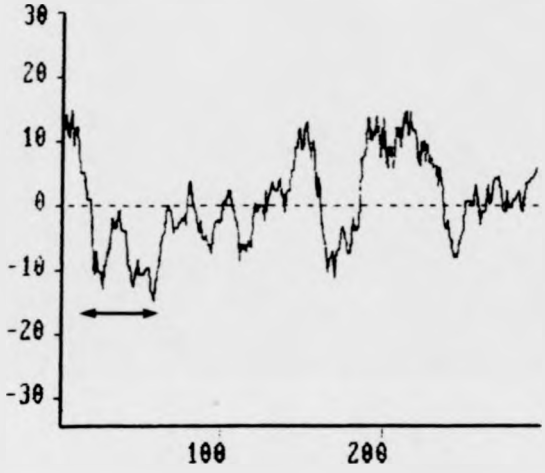


Fig. 4.10 : Hydropathy profiles of PVM strains J3666 and 15 G proteins.

Hydropathy plots of predicted G proteins of a) PVM strains J3666 and b) 15 and c) HRSV (strain A2). The procedure of Hopp and Woods (1981) with a window of 10 amino acids was used. The hydrophobic regions (indicated) at the N-termini are thought to be the transmembrane regions.

presumed single polymerase entry site (Dickens *et al*, 1984, Ling *et al*, 1992) than those of PVM or HRSV, relatively lower amounts of the TRTV G mRNA would be synthesised than for PVM or HRSV virus, thus resulting in the decreased expression of the TRTV G protein. Thus, the presence of an initiation codon upstream of the major ORF's of the G genes of PVM strains 15 and J3666 and HRSV suggest the possibility of down-regulation of gene expression at the translational level as an alternative to the low transcriptional level presumably found for TRTV. No evidence for the synthesis of polypeptides from upstream ORF's of PVM or HRSV have been reported.

As can be seen from Fig. 4.10, the hydropathy plots of PVM strains J3666 and 15 resembles that of HRSV when considering the N termini of these proteins. The hydrophobic transmembrane regions of HRSV and PVM strain J3666 are preceded by a short hydrophilic, cytoplasmically located region which is absent in the hydropathy plot of PVM strain 15. The absence of an apparent cytoplasmic tail in the G protein of PVM strain 15 results from the loss of the cytoplasmic tail present in the G protein of strain J3666 due to the addition of a single A residue (in PVM strain 15 G gene) within the stretch of A residues located at positions 169 to 172 of strain J3666 (Fig. 4.4). The importance of this is unclear, but it is generally believed that cytoplasmic tails of the G and F glycoproteins of the *Paramyxoviruses* interact with the matrix protein during virus maturation (Peeples, 1991; Ray *et al*, 1991). However the first three amino acids of the G protein of PVM strain 15 may protrude into the cytoplasm and could be stabilised by the addition of a lipid chain to the cysteine residue (second amino acid) of this short tail. Veit *et al* (1990, 1991) have shown the M2 protein of influenza virus and the HA protein of influenza B and influenza C viruses to be acylated. The HA protein of influenza virus B was found to contain palmitic acid whereas that of influenza C was modified by the addition of stearic acid (Veit *et al*, 1990). Furthermore, Collins and Mottet (1992) demonstrated the palmitylation of the G protein of HRSV although the precise location for addition of

palmitate was not identified. Arumugham *et al* (1989) also observed the addition of palmitate to the HRSV F protein. However in our hands, immunoprecipitation analysis revealed no incorporation of ^{3}H-palmitic acid into PVM strain 15 G protein (data not shown).

Although virions of a ts mutant of VSV lacked visible G glycoprotein (seen as spike-like projections) at the non-permissive temperature, transmembrane and cytoplasmic remnants of the G protein were found within the viral envelope, suggesting the cytoplasmic tail to be indispensable for efficient virus budding (Schnitzer *et al*, 1979, Metsikko and Simons, 1986). Stricker and Roux (1991) demonstrated that a ts mutant of Sendai virus lacking not only the extracellular domain, but also the transmembrane and cytoplasmic domains of the HN glycoprotein at the non-permissive temperature, was able to bud from infected cells indicating that these regions are dispensable for virus maturation. They concluded that in VSV, which has only one major glycoprotein, the transmembrane and cytoplasmic tail of the G protein were both essential for virus budding whereas in Sendai virus which possesses two major surface glycoproteins (F and G proteins) only one of these proteins was required for efficient virus budding. Alternatively, a cytoplasmically located domain could be added to the N terminus of the G protein of PVM strain 15 by either the insertion of non-templated bases or the deletion of a single residue during viral transcription.

Ling and Pringle (1989b) observed two species of virion associated G proteins in immunoprecipitations of supernatant from PVM strain 15 infected cells. Since the initiation codon for the major polypeptide of the G gene of strain 15 is in a weaker context for initiation than that found upstream, then according to the ribosome scanning model (Kozak, 1989), a majority of ribosomes would translate from the initiation codon for the first ORF. Additionally, since the termination codon for this ORF is located downstream of AUG_{183}, which initiates synthesis of the major

polypeptide, these ribosomes would bypass initiation at this AUG. Initiation of polypeptide synthesis at the AUG_{246} located in frame and downstream of AUG_{183} would result in a polypeptide that would be secreted from infected cells and which should not be associated with the virions. Thus, it remains possible that by some mechanism, following initiation of polypeptide synthesis at the AUG_{83}, the reading frame switches to that of the major open reading frame. However, sequence analysis between the AUG_{83} and AUG_{183} reveals the presence of a termination codon (UAA; at position 168-170 of the sequence shown for strain 15 in Fig. 4.4) upstream and in frame with the AUG_{183}. Thus any modifications that would lead to the switching of reading frames from that of the first ORF to that of the major polypeptide should be found between this termination codon and the termination codon for the small ORF located downstream of AUG_{183} positions 188 to 190 (Fig. 4.4).

Initiation of polypeptide synthesis has been shown to occur at non AUG codons. The C' proteins of human PIV1 isolates have been shown to initiate at GUG codons (Boeck *et al*, 1992, Power *et al*, 1992) whereas that for Sendai virus is known to initiate at an ACG codon (Curran and Kolakofsky, 1988, Gupta and Patwardhan, 1988). Analysis of the PVM G gene sequence would suggest that such an option is unlikely for PVM since initiation of a polypeptide from a non AUG codon from within the G gene of PVM strain 15 would not lead to the addition of a cytoplasmic tail because of the presence of a termination codon located just upstream and in frame with the initiation codon for the major polypeptide. Thus, if any polypeptide were to be initiated from a non AUG codon it would be terminated before translation of the major polypeptide sequence began. Insertion of non-templated bases resulting in a shift of reading frame during transcription of the P/V genes of SV5 (Thomas *et al*, 1988), measles virus (Cattaneo *et al*, 1989), Sendai virus (Vidal *et al*, 1990) and human parainfluenza type 2 virus (Ohgimoto *et al*, 1990), mumps virus (Paterson and Lamb, 1990) and the P/D genes of HPIV3 (Galinski *et al*, 1992) has

been observed. These viruses can be divide into two group based on the resulting polypeptides following insertion of non-templated bases. In measles virus and Sendai virus the P gene which spans most of the length of the gene is transcribed from the major ORF whereas the V proteins are accessed following insertion of non-templated bases. For SV5, mumps virus and PIV2 the V proteins are synthesised from the first AUG and insertion of non-templated bases results in access to the P proteins. In all cases insertion of non-templated G residues occurs at a run of G residues located within the P genes. Vidal *et al* (1990) demonstrated this type of transcriptional modification to be a function of the virus encoded polymerase. Sequence analysis of mRNA isolated from cells expressing the P gene of Sendai virus following infection with a recombinant vaccinia virus revealed the absence of any mRNA species in which non-templated bases had been added. However sequence analysis of mRNA of the same cell line infected with Sendai virus contained species with insertion of non-templated G residues (Vidal *et al*, 1990). These results suggest the possibility that such an event might also occur during transcription of PVM strain 15 G gene to generate a polypeptide with a longer N-terminus region. The presence of a stop codon (UAA; at position 168-170 of the sequence shown for strain 15 in Fig. 4.4) upstream and in-frame with the major ORF of the G gene of PVM strain 15, indicates that insertion of non-templated bases could only occur between positions 168 and 180, most probably with the run of A residues of the sequence shown in Fig. 4.4. Although non-templated insertion of bases within the P genes of *Paramyxoviruses* is thought to occur by stuttering of the viral polymerase at a run of G residues by a mechanism similar to that of polyadenylation, it would be possible for insertion of non-templated bases to occur within the run of A residues at positions 169 and 173 of the nucleotide sequence of the G gene of PVM strain 15. Nucleotide sequences of seventeen independent M13 mp18 phage clones containing DNA amplified by PCR representing the sequence between positions 83 and 180 of PVM strain 15 were found to be identical to that obtained by

sequence analysis of PVM strain 15 (Fig. 4.6). The results suggests insertion of non-templated bases during G gene transcription of PVM strain 15 is unlikely to occur at this position. One speculative possibility is that there could be frameshifting due to pseudoknot formation between the first and second ORF, as found during translation for various retroviruses such as Rous sarcoma virus (Jacks and Varmus, 1985), mouse mammary tumour virus (Moore *et al,* 1987; Jacks *et al,* 1987) and the avian coronavirus, infectious bronchitis virus (Brierley *et al,* 1989). Such a phenomenon could alter the reading frame of the first ORF to that of the second ORF of the G gene of PVM strain 15 resulting in the addition of amino acid residues to the N-terminus of the polypeptide synthesised from the second ORF.

However, the possibility exists that a secretory form of the G protein could have been precipitated from PVM infected cell supernatant, which would then be interpreted as virion associated. Two species of the HRSV G protein have also been observed although one of these is secreted into the medium (Hendricks, 1987). Hendricks *et al* (1988) suggested the soluble form was a result of cleavage within the transmembrane region of the HRSV G protein. However, Roberts *et al* (1993) presented evidence that this secreted form was a result of initiation at an internal AUG located downstream (or within) of the transmembrane region producing a protein. which is proteolytically digested to a mature form. It is possible that the PVM G protein could be cleaved by signal peptidases at the sequence Ser-Asn-His-Lys found downstream of the internal methionine, thus releasing a soluble form of the protein. This cleavage sequence is very similar to the one found in the G protein of HRSV (Ala-Asn-His-Lys; Hendricks *et al,* 1988).

CHAPTER 5

CELLS SURFACE EXPRESSION OF PVM G PROTEINS

5.1. INTRODUCTION :

In the previous chapter the determination of the nucleotide sequence of the G genes of PVM strains J3666 and 15 was described. Only two amino acid substitutions were found within the extracellular region of the predicted proteins. However, several differences were found within the N-terminus coding region of the G gene of PVM strain 15 when compared to that of PVM strain J3666. Most notable was the absence of a hydrophilic domain at the N terminus of the G gene of PVM strain 15 which presumably would form the cytoplasmically located domain with strain J3666. Additionally, the initiation codon (AUG_{183}) of the major polypeptide of the G gene of strain 15 was in a less favourable context for initiation than that for the shorter ORF located upstream of the major ORF. Thus, expression from the major ORF is likely to be diminished or reduced by the presence of this upstream AUG. The aim of the work described in this chapter was to determine whether expression from AUG_{183} resulted in the incorporation of the synthesised polypeptide into the cell membrane. Since the presence of the first AUG_{83} might down-regulate expression of the major ORF, it was deemed necessary to remove this AUG for subsequent expression work.

5.2. RESULTS :

5.2.1. Removal of first AUG :

Fig 5.1 shows a diagrammatic representation of the procedure adopted for the removal of the first initiation codon (AUG_{83}) of the G gene of PVM strain 15. The 5' end of a clone, designated G(3), representing the G gene of PVM strain 15 was digested with exonuclease *Bal31* following linearisation with *EcoRI*. At regular intervals aliquots were removed and the reaction stopped by the addition of EDTA. Truncated fragment ends were end-filled with Klenow DNA polymerase and the gene was released with digestion with *PstI*. The truncated fragments were purified from a low melting point agarose gel (Fig. 5.2) and ligated into *SmaI* and *PstI* digested

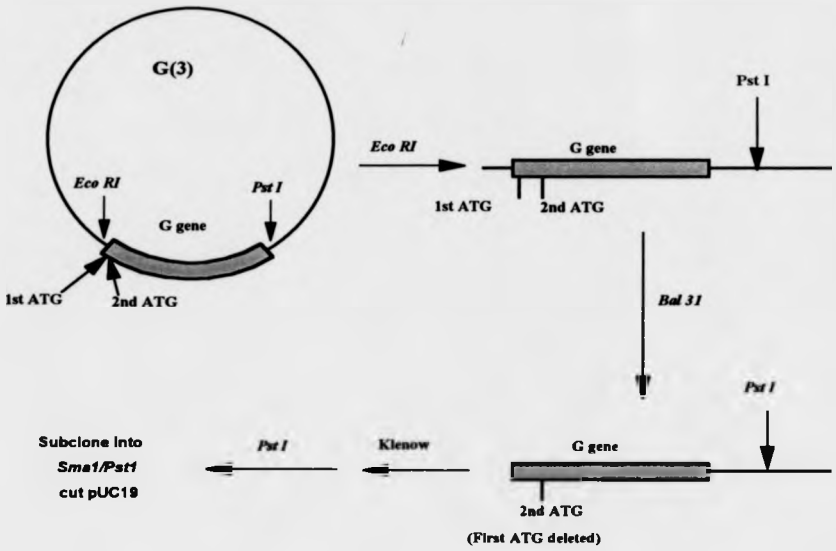


Fig. 5.1 : Removal of first available AUG of PVM strain 15 G gene.

Diagrammatic representation of the procedure used in the removal of the first ATG located of the G gene of PVM strain 15 and the subsequent cloning of truncated inserts.

pUC13 DNA. The sizes of these truncated fragments were estimated from miniprep analysis following digestion of miniprep DNA with EcoRI and PstI (not shown). Clones in which the insert size was calculated to be reduced by approximately 100 nucleotides were selected for nucleotide sequence analysis. *EcoRI* and *PstI* digested fragments of miniprep DNA were purified from a low melting point agarose gel and ligated into *EcoRI* and *PstI* digested M13 mp19 DNA. One of the resulting clones, designated 5.11, represented the nucleotide sequence starting at position 148 of the G gene of PVM strain 15 and was selected for expression work (Fig. 5.3).

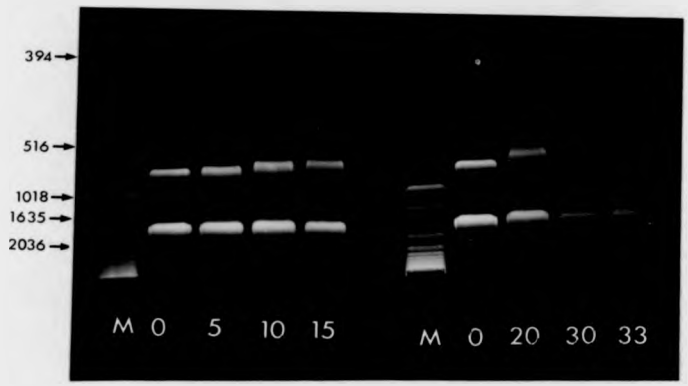


Fig. 5.2 : 5' end truncations of PVM strain 15 G gene.

Low melting point agarose gel stained with ethidium bromide showing fragments of the G gene of PVM strain 15 following linearisation with *EcoRI*, digestion with exonuclease *Bal31*, end-filled with Klenow DNA polymerase and released with digestion with the *Pst1*. Numbers shown below each lane represent intervals (in minutes) at which aliquots of the *Bal31* reaction were stopped by the addition of EDTA. M represents the DNA marker lane.

5.2.2. Transient expression of full-length and truncated PVM G gene :

5.2.2.1. CONSTRUCTION OF TRANSIENT EXPRESSION DNA :

Clones representing full-length (G(3)) and truncated (5.11) PVM strain 15 G gene were subcloned into the SV40 based transient expression vector, pSVL (Fig. 5.4). 5.11 plasmid DNA was digested with *EcoRI* and *XbaI*, end-filled with

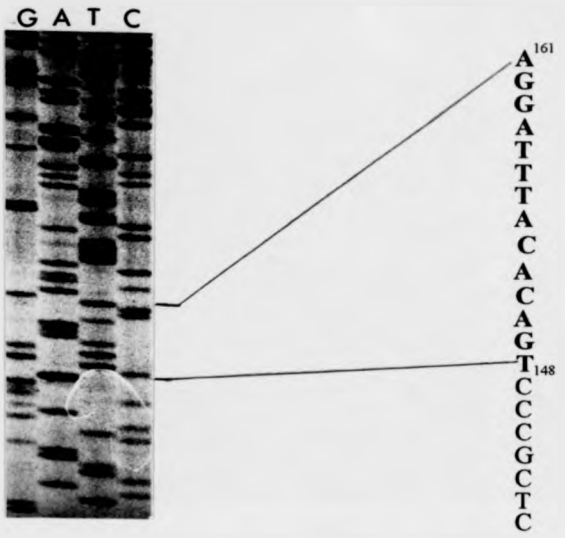


Fig. 5.3 : 5' nucleotide sequence of PVM strain 15 G gene clone lacking first available AUG.

Autoradiograph showing the nucleotide sequence at the 5' end of PVM G gene clone 5.11. The sequence was found to contain nucleotide sequence located after nucleotide 148 (indicated) of the sequence for the full-length PVM G strain 15 shown in Fig. 4.4. PVM strain 15 G gene specific nucleotide sequences are shown in bold.

Klenow DNA polymerase and the insert purified from low melting point agarose gel. The insert was then ligated into *SmaI* digested pSVL. Clones representing the 5.11 insert in i) correct orientation for expression (designated pSVG) and ii) incorrect orientation for expression (designated pSVGR) were identified following digestion of miniprep DNA with restriction endonucleases *HindIII* or *HincII* (Fig. 5.5).

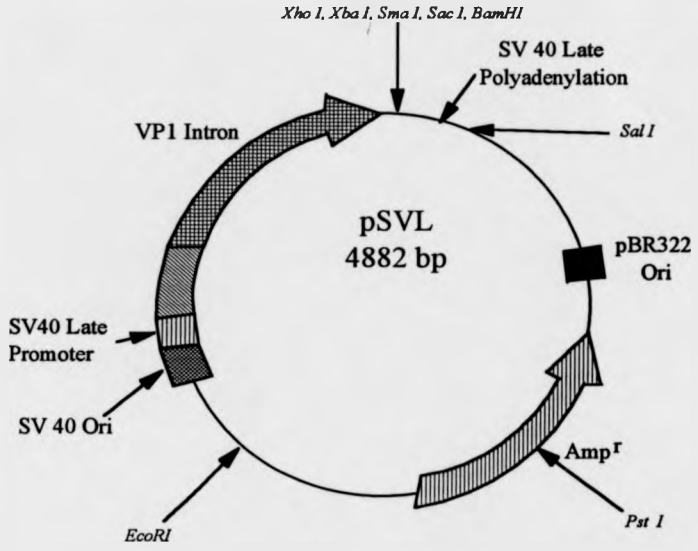


Fig. 5.4 : pSVL, a SV40 based expression vector.

SV40 based shuttle vector used for expression of PVM strain 15 G gene and the truncated form of this gene (clone 5.11). The genes were cloned into the multiple cloning site and expression directed by the SV40 late promoter as described in the text. Polyadenylation of mRNA occurs following correct splicing of transcripts using the SV40 VP1 processing signals.

Plasmid DNA of clone G(3) was digested with *XbaI* to release a fragment containing the whole of the PVM G gene and this was purified from a low melting point agarose gel and ligated into *XbaI* linearised pSVL. Clones representing the full-length G gene in i) the correct orientation for expression (designated pSG(3)) and ii) the incorrect orientation for expression (designated pSGR(3)) were selected following digestion of miniprep DNA with restriction endonucleases *HindIII* or *HincII* (Fig. 5.5).

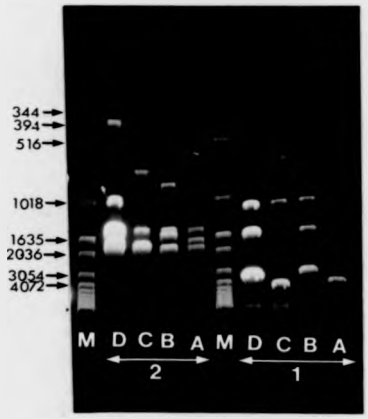


Fig. 5.5 : Restriction endonuclease analysis of pSVL-G gene constructs.

1% agarose gel stained with ethidium bromide showing 1) *HindIII* and 2) *HincII* restriction endonuclease digestion of a) pSVG, b) pSVGR, c) pSCG(3) and d) pSCGR(3) pSVL vector constructs (see text for descriptions). M represents the DNA marker lane.

5.2.2.2. IN VITRO EXPRESSION OF PSG(3), PSGR(3) PSVG AND PSVGR :

CsCl density gradient purified DNA of pSVG, pSVGR, pSG(3) or pSGR(3) was introduced into Cos-7 or L cells using a variation of the Chen and Okayama (1987) transfection protocol as described in the chapter 2, section 2.6.2.1. 48 hours post-transfection, cells were seeded onto multi-spot slides and incubated overnight at 37°C. Expression of the G gene was visualised by indirect immunofluorescence using a G monoclonal antibody specific for PVM strain 15 G protein (19/1/C9; Ling and Pringle, 1989b) and FITC conjugated goat anti-mouse antibody. A very low percentage of transfected cells was found to express the full-

length G gene and that from plasmid 5.11, although the intensities of expressing cells appeared to be equivalent. Fig. 5.6 shows indirect immunofluorescence of acetone fixed transfected Cos-7 and L cells. The results indicate that the proteins produced from the both the full-length and truncated G gene of PVM strain 15 are located in the cytoplasm. No fluorescence was observed for either pSVGR, pSGR(3) or mock transfected cells (not shown). It was not possible to demonstrate surface expression of transfected cells, since unfixed cells did not adhere to the slides during the indirect immunofluorescence procedure. However, cells found to be expressing either the full length G gene or 5.11 appeared to have an abnormal morphology possibly due to the presence of membrane associated G protein.

5.2.3. Construction of recombinant vaccinia virus :

5.2.3.1. CONSTRUCTION OF SHUTTLE PLASMIDS :

Full length G genes of PVM strain J3666 and strain 15 and 5.11, representing the G gene of strain 15 lacking the first initiation codon, were subcloned into a vaccinia virus shuttle vector, pSC11 (Fig. 5.7, Chakrabarti *et al*, 1985). This vector was chosen to construct recombinant vaccinia virus since expression of foreign DNA is controlled by the vaccinia virus early promoter $P_{7.5}$. Additionally the cloning site for this plasmid is flanked by segments of the vaccinia virus thymidine kinase (TK) gene to allow homologous recombination into super-infecting virus genomic DNA, and a copy of the *E. coli* lacZ gene used to identify true recombinant vaccinia virus. Homologous recombination (Fig. 5.8) occurs following transfection of wild type vaccinia virus infected cells with recombinant vaccinia virus shuttle plasmid. Since homologous recombination occurs within the TK gene DNA sequence and segments of the TK gene found within the shuttle vector, insertion of foreign DNA results in the inactivation of the TK gene in recombinant vaccinia virus to give a TK^- phenotype.

a)

b)

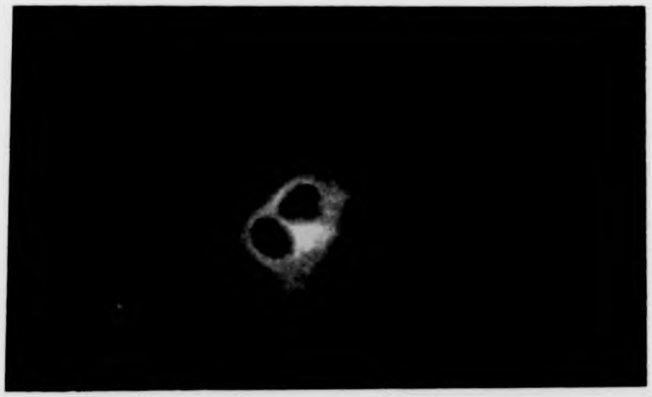


Fig. 5.6 : Immunofluorescence of cells expressing PVM strain 15 G gene.

Immunofluorescence staining of Cos-7 cells transfected with either a) full-length PVM strain 15 G gene or b) construct 5.11 expressed from a SV40 based expression vector, pSVL (see text for details).

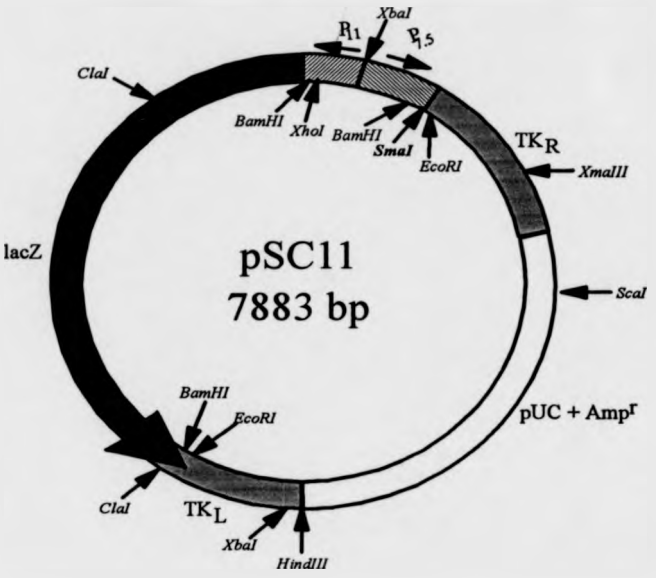


Fig. 5.7 : pSC11, a vaccinia virus shuttle vector.

Vaccinia virus shuttle vector carrying the *E. coli* lacZ gene and segments of the vaccinia virus thymidine kinase (TK) gene used in construction of recombinant vaccinia virus. PVM G genes of strains J3666 and 15 and the 5.11 insert were cloned into the unique *SmaI* of this vector.

Recombinant viruses are then selected for in HuTK$^-$ 143 cells in the presence of 5-bromo-2'-deoxyuridine (BrdU). Lethal incorporation of BrdU into the virus genome occurs only after phosphorylation of BrdU by thymidine kinase expressed by wild type vaccinia virus. True recombinant vaccinia virus is then distinguishable from spontaneous wild type revertants by their blue appearance when X-gal is added to the overlay medium. Thus, selection of recombinant vaccinia virus was considered to be more likely if pSC11 rather than a shuttle vector carrying the thymidine kinase markers but not the lacZ gene was used.

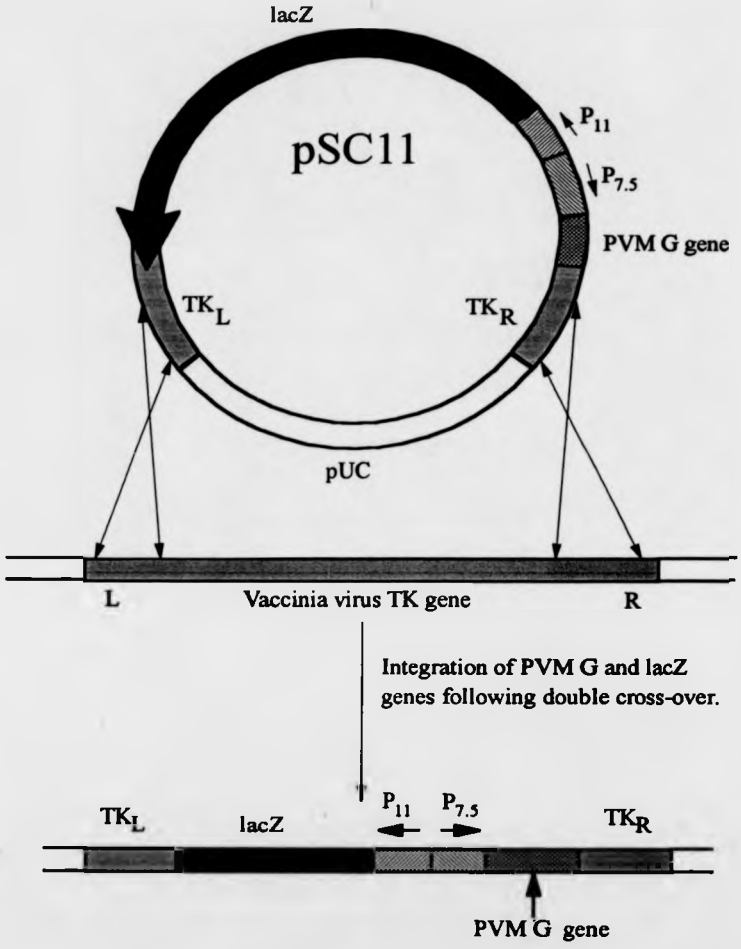


Fig. 5.8 : Homologous recombination between genomic vaccinia virus DNA and pSC11.

Diagrammatic representation of the integration of PVM G genes into the vaccinia virus genome following homologous recombination between the TK gene of wild type vaccinia virus and the TK gene segments in the pSC11 recombinant shuttle DNA in infected-transfected cells.

The full-length G gene of PVM strain J3666 was digested with *EcoRI* and *PstI* whereas that of strain 15, clone G(3), was digested with *XbaI* in order to release the gene inserts from pBS and pUC13 vector DNA respectively. The insert from plasmid 5.11 DNA was released following digestion with *EcoRI* and *XbaI*. Digested DNA of all three samples was phenol/chloroform extracted and ethanol precipitated. The inserts were end-blunted with T4 DNA polymerase, purified from a low melting point agarose gel and ligated into *SmaI* linearised pSC11. The orientation of the inserts within several clones for each construct was determined by digestion of miniprep DNA with either *BamHI* or *HindIII* and *XhoI*. Two clones for each ligation were selected, one representing the insert in the correct orientation for expression whilst the other in the incorrect orientation for expression. Fig. 5.9 shows *BamHI*, and *HindIII* and *XhoI* digested CsCl density gradient purified DNA of the G gene of PVM strain J3666 in the correct orientation for expression, designated pSCG3666, and the incorrect orientation for expression, designated pSCGR3666. Similarly, constructs containing G(3) or 5.11-derived inserts in the correct orientation for expression were designated pSCG(3) and pSCG respectively whereas those in the incorrect orientation were designated pSCGR(3) and pSCGR respectively.

5.2.3.2. SELECTION OF RECOMBINANT VACCINIA VIRUS :

HeLa cells were infected with a wild type vaccinia virus at a multiplicity of infection (m.o.i.) of 2 pfu/ml. At 1 hour post-infection, CsCl-density gradient purified DNA of pSCG, pSCGR, pSCG(3), pSCGR(3), pSCG3666 or pSCGR3666 were introduced into these infected cells using either of the transfection protocols described in chapter 2, section 2.6.2. Transfected-infected cells were harvested once all the cells had become detached from their support.

Aliquots of supernatants containing mixture of wild-type and recombinant vaccinia virus were serially diluted and used to infect monolayers of

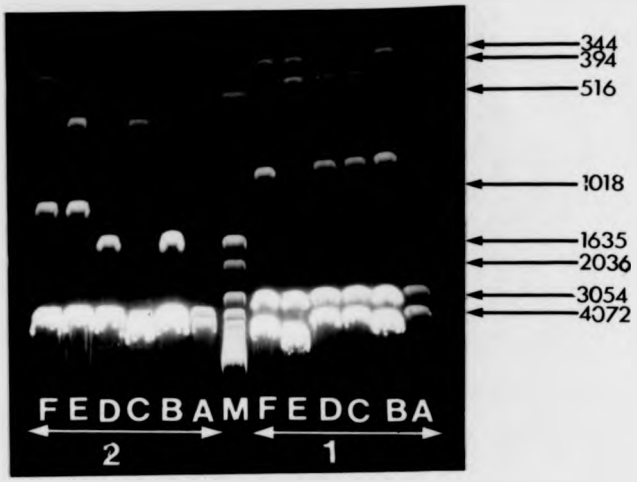


Fig. 5.9 : Restriction endonuclease analysis of recombinant vaccinia virus shuttle vector.

1% agarose gel stained with ethidium bromide showing 1) *HindIII* and *XhoI* digest and 2) *BamHI* restriction endonuclease digests of a) pSCG, b) pSCGR, c) pSCG(3), d) pSCGR(3), e) pSCG3666, and f) pSCGR3666 representing recombinant vaccinia virus shuttle vector containing the G genes of both strains of PVM (see text for detailed descriptions). M represents the DNA marker lane.

HuTK$^-$ 143 cells grown in 6-well tissue culture plates. Following adsorption of virus, cells were overlaid with a 1% (w/v) agar overlay containing GMEM supplemented with FCS (final concentration of 2%) and 25 µg/ml BrdU and incubated for 48 hours at 37°C, 5% CO_2. To distinguish recombinant vaccinia virus from spontaneous revertants, 1% (w/v) of agar containing 20 µg/ml X-gal was added to the infected cells and incubated overnight at 37°C, 5%. Fig. 5.10 shows blue colouration of recombinant vaccinia virus during the selection process. Visible individual blue plaques were transferred into 0.5 ml of GMEM supplemented with FCS to a final concentration of 5%.

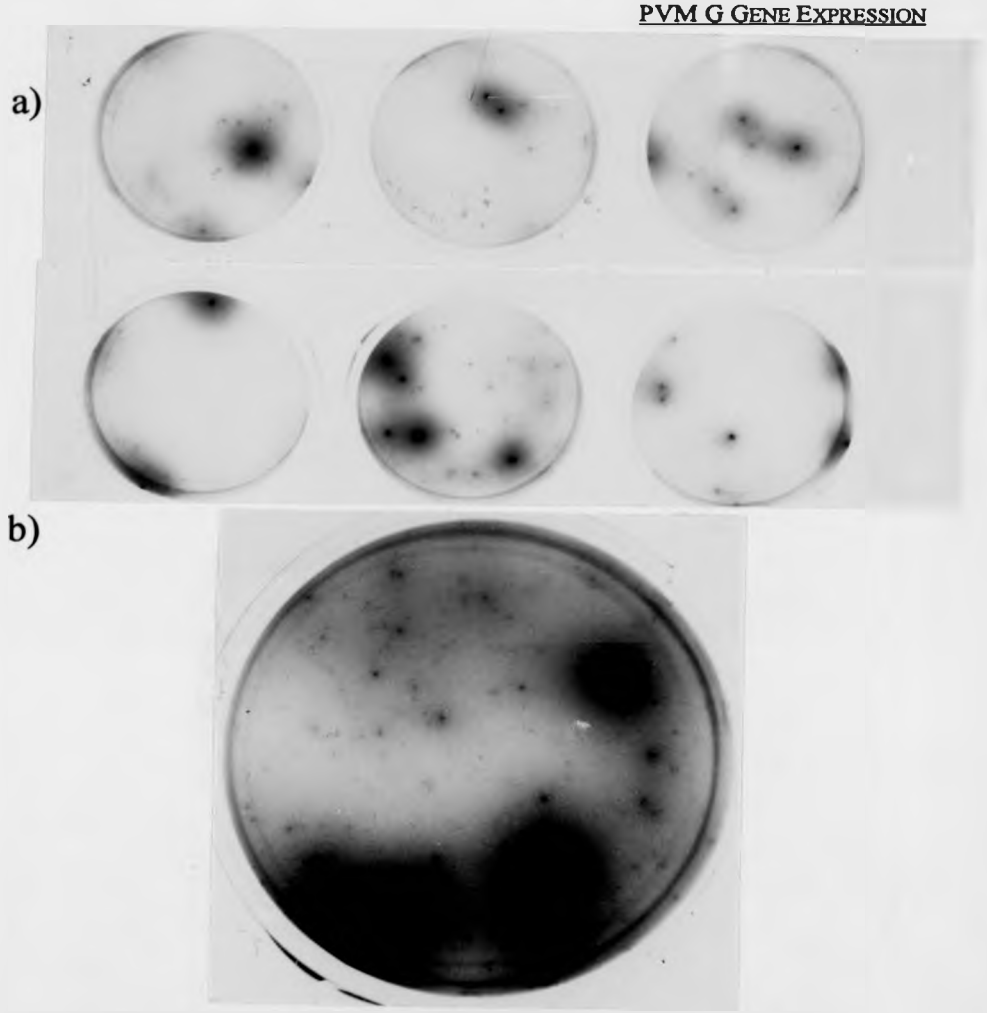


Fig. 5.10 : Visualisation of recombinant vaccinia virus.

Photograph showing expression of the lacZ gene in TK^- cells, of a) six-well plate and b) close up of one of the wells, of recombinant vaccinia virus plaques (blue) containing full-length PVM strain 15 G gene following infection with virus extracted from BS-C-1 cells that had been infected with wild type vaccinia virus and transfected with construct pSCG(3) DNA.

Several recombinant plaques (blue) from each construct were transferred into GMEM supplemented with FCS (final concentration of 2%) and underwent three cycles freeze-thaw. The plaque purification process involved amplification of recombinant vaccinia virus in tissue culture cells which would then undergo plaque purification. Aliquots of each recombinant vaccinia virus were added to monolayers of HuTK$^-$ 143 and incubated at 37°C, 5% CO_2 in the presence of GMEM supplemented with FCS (2% final concentration) and containing 25 μg/ml BrdU. Unfortunately, no replication of these recombinant plaques was observed as measured by their ability to cause CPE. Similar results were obtained when recombinant plaques were amplified in HeLa cells. Attempts to isolate replicative recombinant vaccinia virus from overlays containing low melting point agarose gel or by sonication of selected plaques also proved fruitless.

5.2.4. Expression of PVM G genes directed by T7 RNA polymerase recombinant vaccinia virus :

The failure to construct recombinant vaccinia virus carrying the PVM G genes necessitated the adoption of an alternative strategy for demonstrating the surface location of the G proteins of both strains of PVM. However, expression of genes have been achieved by transfection of cells infected with recombinant vaccinia virus, vTF7-3 (Fuerst *et al*, 1986) expressing the T7 RNA polymerase gene, following transfection with plasmid DNA in which expression of genes are directed from the T7 promoter. Thus, it was decided to express the G genes of both PVM strains using this system.

5.2.4.1. CONSTRUCTION OF PGEM1 RECOMBINANT PLASMIDS :

The G gene of PVM strain J3666 was released from pBS vector DNA following restriction endonuclease digestion with *EcoRI* and *PstI* whereas that of strain 15, clone G(3) was released from pUC13 by digestion with *XbaI*. Plasmid 5.11 DNA was digested with *EcoRI* and *XbaI* to release the insert. Following purification of

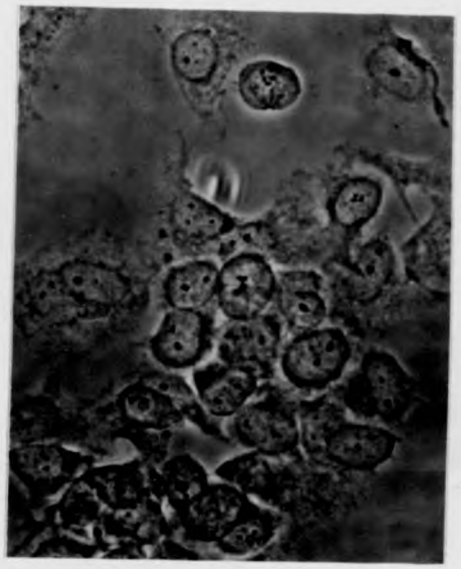
the gene inserts from a low melting point agarose gel, the inserts were ligated into *EcoRI* and *PstI*, *XbaI* and *EcoRI/XbaI* digested pGEM1 expression vector DNA respectively. This placed the transcription of the genes under the control of the T7 RNA promoter. Bacterial colonies containing the appropriate recombinant plasmids were identified by digestion of miniprep DNA with *EcoRI* and *PstI*. Constructs of pGEM1 and the G gene of strain J3666 and construct 5.11 were designated pGEM1-G3666 and pGEM1-5.11. Clones of pGEM1-G(3), constructed by ligation of *XbaI* digested pGEM1 and G(3), that would result in transcriptional control from the T7 RNA promoter were distinguishable following digestion of miniprep DNA with *HindIII*, since a single *HindIII* restriction site is located 3' of AUG_{183} within the G genes of both strains of PVM and the other *HindIII* restriction site located within pGEM1 vector (not shown).

5.2.4.2. EXPRESSION OF PVM G GENES :

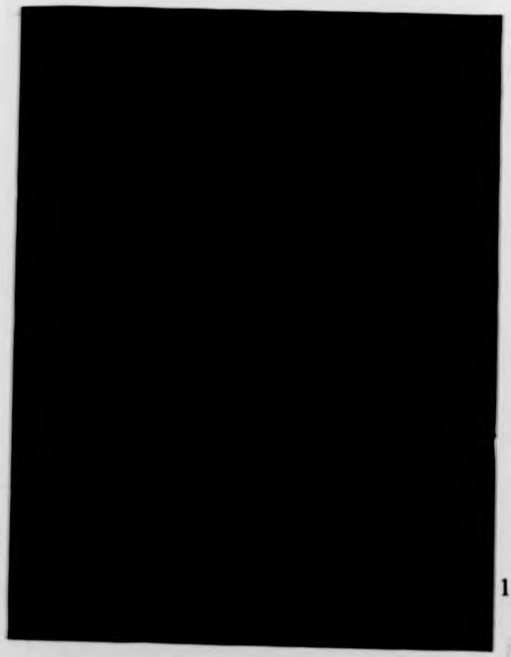
BS-C-1 cells grown on glass coverslips overnight were infected with recombinant vaccinia virus carrying the T7 RNA polymerase gene at a multiplicity of infection of 5 pfu/ml. Following adsorption of the virus for 1 hour at 37°C, TransfectACE-DNA complex diluted in Opti-MEM reduced medium, containing the appropriate CsCl-density gradient purified DNA was added to the infected cells and incubated for 6 hours at 37°C in a 5% CO_2 atmosphere. GMEM containing 20% FCS was added to give a final concentration of 10% and the cells incubated for 15-18 hours. Expression of PVM G protein was visualised by indirect immunofluorescence following staining with Texas-red as described in chapter 2, section 2.7.10. Transfected/infected cells showing surface expression were fixed with 2% paraformaldehyde only after incubation with streptavidin-Texas red conjugate during the immunofluorescence staining procedure following incubation of the cells with MAb 19\1\C9 (Ling and Pringle, 1989b). Fig. 5.11 shows the surface location of G proteins

a)

b)



c)



d)

e) f)

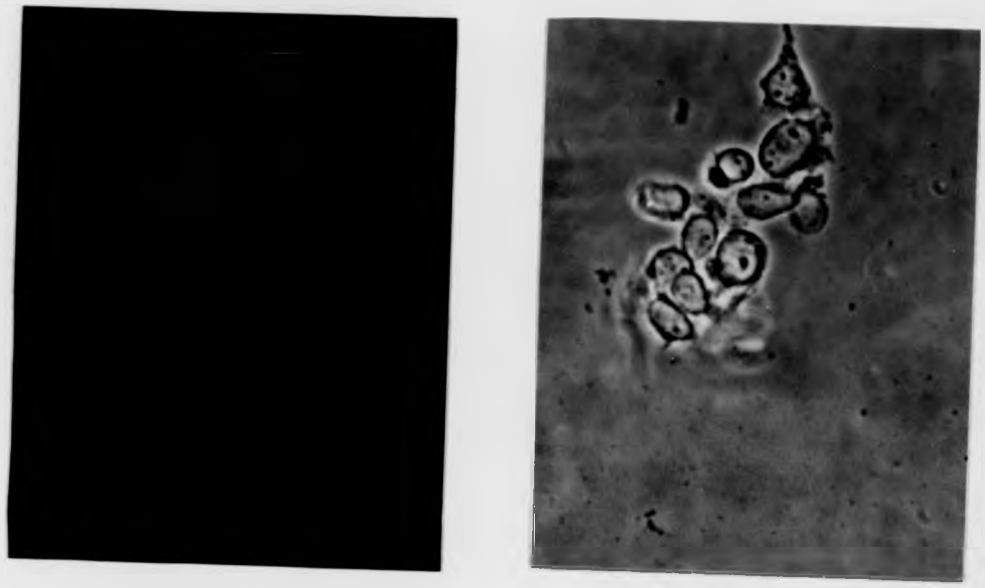


Fig. 5.11 : Cell surface location of PVM G proteins.

Photographs showing surface location of a) PVM strain J3666 G gene c) full-length strain 15 and e) 5.11 inserts expressed under the control of the T7 RNA polymerase promoter in BS-C-1 cells infected with recombinant vaccinia virus expressing the T7 RNA polymerase. Phase contrast photographs of the expressing cells are shown for b) PVM strain J3666 G gene, d) full-length strain 15 G gene and f) 5.11 insert. Cells were fixed in 2% paraformaldehyde only after incubation with streptavidin-Texas red conjugate during the immunofluorescence staining procedure

of PVM directed from pGEM1-G3666, pGEM1-5.11 and pGEM1-G(3) recombinant plasmid constructs.

5.3. DISCUSSION :

The expression of the G genes of both strains of PVM in tissue culture cells was visualised by indirect immunofluorescence. Since the predicted G protein of PVM strain 15 lacks the putative cytoplasmic domain, which is thought to stabilise the protein whilst imbedded within the lipid bilayer, it is possible that a cytoplasmic tail could be added following possible ribosomal frameshifting after initiation from the first available AUG. Additionally, it was possible that since AUG_{183} of PVM strain 15 G gene is in a poorer context for initiation than AUG_{83} (Kozak, 1986), expression of the major polypeptide from the full-length clone (G(3)) may be down-regulated by the presence of this upstream AUG. To determine whether the G protein lacking this cytoplasmic tail could be correctly and stably inserted into the lipid bilayer, the first AUG was removed by digestion with exonuclease *Bal31*. Thus, any protein synthesised from this construct would be as a result of initiation from the second AUG codon alone.

Expression of the full-length and truncated form of the G gene of PVM strain 15 in Cos-7 and L cells directed by the SV40 late promoter was visualised by indirect immunofluorescence using MAb 19/1/C9 (Ling and Pringle, 1989b) and confirm the coding assignment of this gene. A very low percentage of transfected cells was found to express the full-length G gene and 5.11, although the intensities of expressing cells appeared to be equivalent. Although expression was observed in transfected cells that had been fixed in cold acetone prior to the immunofluorescence staining procedure, expression in unfixed cells could not be observed as they did not remain attached to the slide during the detection procedure. Furthermore, cells shown to express either of the constructs had an abnormal appearance in which the cell

membrane appeared thickened possibly due to incorporation of the G proteins within the cell membrane. These results although not confirming the surface location of the proteins do suggest association of the G protein lacking the cytoplasmic tail with the cell membrane.

Attempts were made to isolate permanent cell lines expressing the PVM G proteins by co-transfection of cells with these pSVL-G constructs with plasmids carrying neomycin or puromycin resistance genes. Although colonies of transfected cells resistant to either G418 (geneticin) or puromycin were obtained, FACS analysis did not show expression of the G genes of PVM strain 15. Numerous unsuccessful attempts were also made to construct recombinant plasmids in which the G genes would have been under the control of the major cytomegalovirus immediate early promoter or the metallothionine I promoter flanked by retroviral terminal LTR's that would have increased integration of the shuttle DNA into the cell chromosomes. Attempts to construct shuttle vector DNA with a view to construction of recombinant adenovirus or recombinant herpes simplex virus also proved fruitless due to apparent instability of the insert which appears to be a feature of some genes encoding glycoproteins.

Cloning of HRSV G gene into a variety of expression vectors has proved difficult (L. A. Ball, personal communication) although expression has been achieved by construction of recombinant vaccinia virus (Olmsted *et al*, 1986; Elango *et al*, 1986; Stott *et al*, 1986). The presence of nucleotide sequences resembling those of prokaryotic promoters have been found in HRSV G genes (P. Cane, personal communication). Thus, expression from an internal AUG is possible with the resulting product toxic to the prokaryotic cell. Thus, attempts were made to construct recombinant vaccinia virus carrying the PVM G genes. Although recombinant plaques could be identified by their blue appearance in the presence of X-gal when grown in TK^- cells in the presence of BrdU, no subsequent amplification of these plaques could

be achieved. It had been suggested that the larger, intense blue plaques (Fig. 5.10) were cells expressing the lacZ product from the recombinant shuttle vector DNA constructs and that the smaller, much fainter blue plaques, represented true recombinant virus. However, further amplification of these fainter plaques could not be achieved. Expression of the lacZ gene in transfected cells could only be achieved in the presence of vaccinia virus since the lacZ gene is under the control of the vaccinia virus P_{11} promoter. It is still unknown as to why amplification of these putative recombinant vaccinia virus plaques was not possible.

Recently, expression of a number of viral genes has been achieved by transfection of cells infected with recombinant vaccinia virus (vTF7-3) expressing the T7 RNA polymerase with plasmid DNA in which the gene to be expressed is under the control of the T7 promoter and it was decided to express the PVM G genes using this system. Expression of the PVM G genes was achieved in BS-C-1 cells infected with vTF7-3 following transfection with DNA plasmid constructs in which the G genes were under the control of the T7 promoter located within the pGEM1 vector. Expression of the PVM strain J3666, full-length strain 15 G gene and 5.11 insert was demonstrated in these infected-transfected cells by indirect immunofluorescence. Furthermore, expression of all three constructs was demonstrated in unfixed infected-transfected cells showing the surface location of these gene products. The results confirm that at least a significant proportion of the gene product expressed from AUG_{183} (in construct 5.11 which lacks AUG_{83}) which lacks the cytoplasmic tail, is transported to the cell surface following correct processing of the protein. The results indicate that signals required for correct processing of the PVM G protein probably are located within the transmembrane region and that the cytoplasmic N-terminal tail plays no significant role during processing of the protein. The significance of these findings are as yet unknown since it is generally believed that the cytoplasmic tails of the G and

F glycoproteins of the paramyxoviruses interact with the matrix protein during virus maturation (Peeples, 1991; Ray *et al*, 1991).

Schnitzer *et al* (1979) failed to observe spike-like projections, corresponding to the G protein, on the surface of virions of a temperature sensitive mutant of VSV at the non-permissive temperature. However remnants of the transmembrane and cytoplasmic domains of the G protein were found associated with the viral envelope, suggesting the cytoplasmic tail to be indispensable for efficient virus budding (Metsikko and Simons, 1986). Owens and Rose (1993) constructed a chimaeric protein consisting of the cytoplasmic tail of the G protein of VSV linked to the transmembrane and extracellular domains of the HIV-1 envelope glycoprotein. Expression of this construct resulted in rescue with VSV defective in the G protein at the non-permissive temperature. The rescued mutant was able to bind $CD4^+$ cells indicating the surface location of the chimaeric protein. Another construct in which the VSV G protein cytoplasmic domain was attached to the envelope protein of the HIV-1 which had a truncated cytoplasmic domain failed to rescue the VSV mutant. These results indicate that the cytoplasmic domain of the VSV G protein should be in very close proximity of the cell membrane in order for the M protein to interact with the cytoplasmic domain of VSV during virus budding. Additionally, the cytoplasmic domain of the VSV G protein contains sequences recognised by the VSV M protein (Owens and Rose, 1993). It seems unlikely that the 3 amino acid long cytoplasmic domain of the G protein of PVM strain 15 could interact with the PVM M protein. It is possible that the G protein could be stabilised by the addition of fatty acids to the cysteine residue located within the three amino acids that form the cytoplasmic domain. Collier *et al* (1991) demonstrated a decapeptide corresponding to the cytoplasmic domain of the influenza virus haemagglutinin protein to inhibit release of virus particles and infectious virions. Recently, Stricker and Roux (1991) demonstrated a ts mutant of Sendai virus lacking not only the extracellular domain, but

also the transmembrane and cytoplasmic domains of the HN glycoprotein at the non-permissive temperature, was able to bud from infected cells. A MAb raised against the N terminal 20 amino acids of the Sendai virus HN protein failed to react with any HN remnants in Western blot analysis. They concluded that in VSV, which has only one major glycoprotein, the transmembrane and cytoplasmic tail of the G protein were both essential for virus budding whereas in Sendai virus which possesses two major surface glycoproteins (F and G proteins) only one of these proteins was required for efficient virus budding.

Wilson *et al* (1990) investigated the effect of deletion of the cytoplasmic tail of the NDV HN protein. 23 of the 26 amino acids comprising the NDV HN protein cytoplasmic tail were deleted and its effect upon translocation of the HN protein across the membrane *in vitro* was examined. The removal of the 23 amino acids leaves the mutant protein with a cytoplasmic tail of just three amino acids with a single positive charge. The level of the mutant protein inserted into membranes in *in vitro* translation experiments in wheat germ extracts was less than the wild type HN protein although approximately 50% of the mutant protein was found to be inserted into the membranes since it was resistant to extraction with EDTA or alkaline pH. Extraction with EDTA results in removal of peripheral proteins pelleted with the membrane fraction whereas alkaline pH treatment removes proteins that are secreted or are not integrated into the membranes. Both glycosylated and unglycosylated forms of the mutant protein were found to be inserted into the cells membrane although the unglycosylated form was sensitive to trypsin digestion suggesting an inverse orientation of this mutant protein in which the C terminus would be located internally. The wild type HN protein was not observed in this inverse orientation (Wilson *et al*, 1990). The authors speculate that presence of the long hydrophilic cytoplasmic domain of NDV HN protein may prevent spontaneous insertion of the protein allowing interaction of the newly synthesised hydrophobic transmembrane domain with the cell

membrane and preventing the inverse orientation of the wild type protein. Although the mutant HN protein of NDV was inserted into the membranes in both orientations, it is not known whether the G protein of PVM strain 15 is also inserted in both orientations since there is no way of detecting cells expressing the PVM G protein in the inverse orientation in the absence of appropriate MAb's. Thus, cells expressing the PVM G protein in both the correct and inverse orientations would be positive in the immunofluorescence staining procedure. This area would be of future interest.

CHAPTER 6

GEL ELECTROPHORESIS ANALYSIS OF PVM G PROTEINS

6.1. INTRODUCTION :

The surface location of the G proteins of a pathogenic and non-pathogenic strains of PVM was described in the preceding chapter. Although the G protein of strain 15, the non-pathogenic strain, lacked the cytoplasmic tail which presumably acts in stabilising the protein while anchored within the cell membrane, it appeared to be processed and transported correctly to the cell surface following visualisation by indirect immunofluorescence in cells transfected with plasmid DNA constructs carrying the PVM strain 15 G gene. However, it remains to be determined whether two virion associated G protein forms can be observed for PVM strain J3666 as was the case for PVM strain 15 (Ling and Pringle, 1989b).

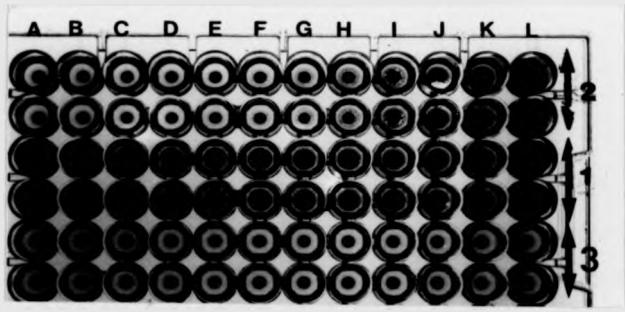
6.2. RESULTS :

6.2.1. Haemagglutination (HA) and inhibition (HI) assays :

PVM differs from HRSV in its ability to agglutinate murine erythrocytes cells. Although it is not known whether the two amino acid substitutions located within the extracellular domain of PVM strain J3666 when compared to strain 15 are as a result of immunological pressures, it can be tested whether these mutations result in a change in haemagglutination by the virus.

Although MAb 19/1/C9 does not neutralise PVM infectivity *in vitro* its ability to inhibit haemagglutination has been previously reported (Ling and Pringle, 1989b). Thus, this MAb was used in haemagglutination inhibition assays to determine reactivity of the MAb to PVM strain J3666. Fig. 6.1 shows the results of HI assays for PVM strains J3666 and 15. Serial dilutions of MAb 19/1/C9 were tested against 5 HA units of PVM strains 15 and J3666 in the presence of 3.2×10^5 mouse red blood cells. The results indicate reactivity of the MAb to be similar in both strains of PVM.

X)



Y)

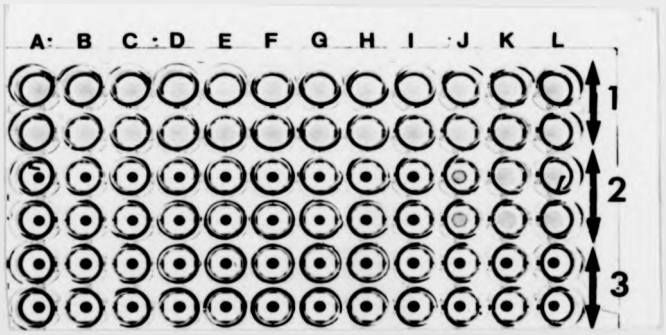


Fig. 6.1 : Haemagglutination inhibition of mouse erythrocytes.

HI assays of both strains of PVM. Virus was added a concentration of 5 HA units. X) HI assay for PVM strain 15. Y) HI assay for PVM strain J3666. A-L represent 2-fold serial dilutions of a 10-fold diluted stock of MAb 19/1/C9 such that A represents a dilution factor of 20. 1) Haemagglutination of mouse red blood cells, 2) haemagglutination inhibition assay and 3) negative control (PBS).

6.2.2. Radioimmunoprecipitations of PVM polypeptides :

Polypeptides of PVM were analysed following immunoprecipitation of ^{35}S-methionine or ^{3}H-glucosamine hydrochloride labelled infected BS-C-1 cells. Unfortunately, immunoprecipitation of infected cells or infected cell supernatants either in the presence of actinomycin D, tunicamycin or monensin proved unsuccessful. Thus, the G protein of PVM strain J3666 in infected cells was not characterised. One reason for the failure to immunoprecipitate the G protein of PVM strain J3666 could be the poor growth of the virus in tissue culture and that it is naturally temperature sensitive with no growth at 37°C (A. Easton, personal communication), although it does not explain the reason for why immunoprecipitation of PVM strain 15 G protein, which has been described by Ling and Pringle (1989b), also failed. One reason could be failure of the MAb (19/1/C9; Ling and Pringle, 1989b) to immunoprecipitate the G protein successfully.

6.2.3. *In vitro* analysis of PVM G proteins :

An alternative approach to the study of the G proteins of both strains of PVM are by *in vitro* transcription and translation analysis.

6.2.3.1. *IN VITRO* TRANSCRIPTION :

The construction of recombinant pGEM1 plasmids in which transcription of full-length G genes of both strains of PVM and a truncated form of the G gene of PVM strain 15 is under the control of the T7 RNA promoter has been described in the preceding chapter. pGEM1 constructs were linearised with restriction endonuclease *SalI* and RNA synthesised *in vitro* using T7 RNA polymerase as described in chapter 2, section 2.7.8. Fig. 6.2 shows the *in vitro* transcribed RNA analysed on a 1% agarose gel stained with ethidium bromide. Although the *in vitro* synthesised RNA appears as a smear possibly due to degradation of RNA, two distinct species of RNA are observed. The first corresponds to the expected size of the *in vitro* transcribed

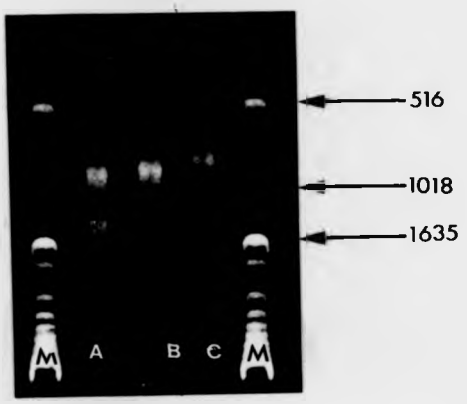


Fig. 6.2 : *In vitro* transcribed RNA.

1% agarose gel stained with ethidium bromide showing *in vitro* transcribed RNA of pGEM1 constructs containing a) full-length PVM strain J3666 G gene, b) full-length PVM strain 15 G gene and insert from construct 5.11. Two species of RNA are prominent for RNA transcribed from strain J3666 although two species were also found for RNA transcribed from the other two constructs. M represents the DNA marker lane.

RNA at approximately 700 base pairs when compared to DNA size markers. The second species appears to be double the size of the first species running at approximately the same size as the insert from which it was transcribed. This species may represent inter-molecular pairing between the smaller species of RNA.

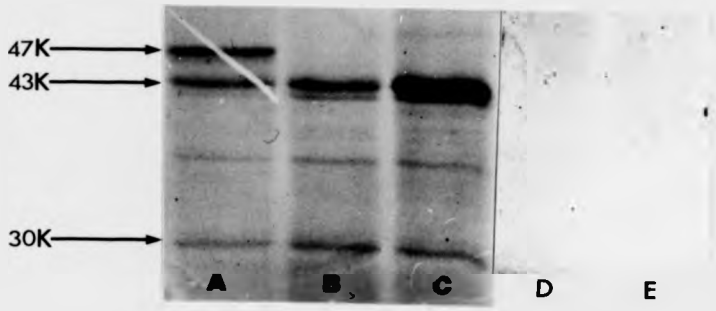


Fig. 6.3 : Autoradiograph of *in vitro* translated products.

Autoradiograph showing *in vitro* translation products from pGEM1 constructs containing a) full-length PVM strain J3666 G gene, b) full-length PVM strain 15 G gene, c) 5.11 insert, d) pGEM1 vector alone and e) no RNA. Cell-free translation was carried out in the presence of wheat germ extract and the polypeptides analysed by SDS-PAGE electrophoresis on a 10% polyacrylamide gel. No wheat germ extract specific polypeptides were observed (lanes d and e).

6.2.3.2. *IN VITRO* TRANSLATION :

In vitro transcribed RNA was used to synthesise proteins *in vitro* using wheat germ extract as described in chapter 2, section 2.7.9. Reactions were carried out for 1 hour at 27°C for wheat germ extract reactions. The products were analysed by SDS-PAGE gel electrophoresis. Fig. 6.3 shows an autoradiograph of the *in vitro* translation products The primary translation products from PVM strain 15 and strain J3666 G genes, and 5.11 insert have a slightly lower mobility than their predicted unglycosylated molecular weights. For 5.11 and PVM strain 15 G gene, the

unglycosylated G protein is approximately 43K whereas that of strain J3666 is 47K. These sizes are approximately 3K larger than their predicted molecular weights. Interestingly, although similar amounts of RNA transcribed from 5.11 and strain 15 G gene were used in the translation system, expression form the 5.11 insert is considerably greater than that of the full-length strain 15 G gene. Also, although the primary translation product of strain J3666 G gene is larger than that of either insert 5.11 or strain 15 G gene, a translation product similar to the size of the primary products 5.11 or strain 15 G gene was clearly observed. Furthermore, several smaller products are seen in the wheat germ translated system.

6.3. DISCUSSION :

The failure to identify immunoprecipitated G protein form either PVM strain 15 or strain J3666 infected cells does not determine whether two forms of the G protein, that were observed for PVM strain 15 (Ling and Pringle, 1989b), are also present in PVM strain J3666 infected cells. An alternative approach to study the number of G gene specific polypeptides synthesised from both strains of PVM was in the use of *in vitro* translation systems and translation of *in vitro* transcribed RNA in the presence of wheat germ extract showed good expression of the G gene products.

As expected, the major polypeptide species from translation of RNA synthesised from either full-length G gene of PVM strain 15 or from insert from construct 5.11 have similar molecular weights of approximately 43K. These are in good agreement with the predicted molecular weight of the unglycosylated PVM strain 15 G protein estimated from SDS-PAGE analysis of immunoprecipitated G protein from PVM infected cells in the presence of glycosylation inhibitors (Ling and Pringle, 1989b). The difference between full-length strain 15 G gene and construct 5.11 is that construct 5.11 lacks the first initiation codon (AUG_{83}) following digestion of the 5' end of PVM strain 15 G gene with exonuclease *Bal31* (Chapter 5, section 5.2.1).

Although similar amounts of *in vitro* transcribed RNA representing these two constructs were used in the wheat germ extract translation system, following estimation of RNA on a 1% agarose gel, translation of construct 5.11 was greater than that of the full-length PVM strain 15 G gene. These results can be accounted for by the ribosome scanning model for translation (Kozak, 1986). During translation of the full-length PVM strain 15 G gene, most of the ribosomes would initiate translation from the first available AUG codon which codes for a small polypeptide of 35 amino acids. Since the termination codon for this frame is located downstream of the initiation codon for the major ORF, at AUG_{83}, only ribosomes that fail to start translation at the AUG_{83} will initiate translation from the second ORF which codes for the major polypeptide. Thus, the presence of an AUG upstream which is in a better context for initiation than that of the major ORF of PVM strain 15 acts in down-regulation of expression of the G protein.

As expected the primary translation product following *in vitro* translation of RNA representing the G gene of PVM strain J3666 has a larger apparent molecular weight than that of PVM strain 15. This is reconciled with the fact that the G protein of PVM strain J3666 is 33 amino acids longer that of PVM strain 15. The G protein of PVM strain 15 appears to lack the amino acids sequence that forms the G protein cytoplasmic tail in PVM strain J3666. The observed molecular weight of 47K is somewhat larger than the predicted unglycosylated molecular weight of 43,584. Additionally, a polypeptide of similar size to the primary translation product of the G gene of PVM strain 15 is also observed following translation of the PVM strain J3666 G gene RNA. Thus, it appears that internal initiation from AUG_{83} (in strain J3666) would result in expression of a polypeptide similar to that of PVM strain 15, which also lacks its cytoplasmically located domain. It is not known whether the polypeptide synthesised from AUG_{182} (in strain J3666) is present in PVM strain J3666 infected cells, since attempts to immunoprecipitate G gene specific polypeptides from PVM

strain J3666 infected cells were unsuccessful. Interestingly, the amounts of each polypeptide synthesised from either AUG_{83} or AUG_{182} appear to be similar even though AUG_{29} is located upstream and out of frame with AUG_{83}. Thus, it appears that the presence of the upstream AUG does not inhibit significantly translation from an internal initiation site at least in the wheat germ system. Similar results were observed by Alkhatib *et al* (1988) were similar amounts of the C polypeptide were observed when full-length P/C gene and the C gene of measles virus were expressed from recombinant human adenoviruses. Here, an out of frame upstream initiation codon, coding for the P polypeptide, did not inhibit synthesis of the C polypeptide from an internal initiation codon. The authors suggest that the ribosomes may be binding directly to a site located internally within the mRNA or near the initiation codon for the C polypeptide. The results obtained for *in vitro* translation of PVM strain J3666 G gene contrast with those for strain 15 in which removal of the upstream initiation codon (AUG_{83}) results in increased expression of the primary translation products from the G gene of PVM strain 15. One drawback using the wheat germ extract translation system is the prevalence of premature terminations. Thus, it is not known whether the lower molecular weight polypeptides for both strains of PVM are as a result of internal initiation or premature terminations.

The presence of two amino acid changes within the extracellular domain of PVM strain 15 G protein when compared to that of PVM strain J3666 may represent in antigenic variation. Although both strains of PVM reacted with MAb 19/1/C9 (see chapter 5), it is not known whether these changes are antigenically important since MAb 19/1/C9 does not neutralise PVM strain 15 infection *in vivo* (Ling and Pringle, 1989b). However, it is possible that either of these changes may be located within the heamagglutinin region of the G gene. Whilst the HA assay confirms the ability of both strains of PVM to agglutinate mouse erythrocytes, results obtained from HI assays confirm the ability of MAb 19/1/C9 to inhibit HA activity of both

strains of PVM. Thus, the contribution of these extracellular differences have yet to be determined.

CHAPTER 7

RIBOSOMAL FRAMESHIFTING IN PVM STRAIN 15 G GENE

7.1. INTRODUCTION :

The failure to identify the G polypeptide species of PVM strain J3666 *in vivo* leaves open the question as to whether the two species of the G protein identified for PVM strain 15 by Ling and Pringle (1989b) are also synthesised by PVM strain J3666. It is possible that the two species of G protein seen in PVM strain 15 infected cells are generated by translation of the major ORF and also by the generation of a larger species by fusing the two ORF's together. Addition of a cytoplasmic tail to the G protein via insertion of non-templated bases or as a result of an error in determining the correct sequence of the 5' end of the G gene of strain 15 have been discounted as discussed in chapter 4. An alternative possibility that could result in the addition of such a cytoplasmically located domain is by ribosomal frameshifting.

The phenomenon of ribosomal frameshifting has been described for a number of viruses and has been implicated in expression of polypeptides from overlapping reading frames. Such a frameshifting event appears to be particularly common in a number of retroviruses. The retroviruses can be divided into three groups depending on their gene order and how the *pol* genes is accessed. Fig. 7.1 shows the partial gene order of some retroviruses as examples. The retroviral *pol* genes can be accessed in one of three ways. A read-through mechanism can occur in which the in-frame *gag* and *pol* genes are separated by a stop codon and the 5' end of the *pol* gene that lacks its own initiation codon may or may not overlap with the 3' end of the *gag* gene. Alternatively, the *gag* and *pol* genes may be separated by the protease encoding *pro* gene, which overlaps both the *gag* and *pol* genes (Jacks *et al*, 1988). When overlapping ORF's are present access to either the *pro* or *pol* genes occurs by ribosomal frameshifting resulting in a change of the reading frame in a -1 direction.

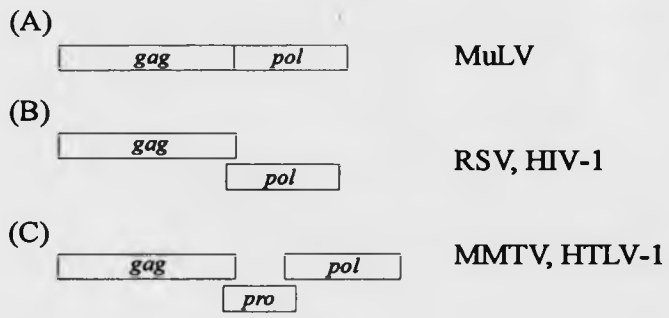


Fig. 7.1 : Partial gene order of retroviruses.

Partial gene order of retroviruses showing the arrangement of retrovirus *gag*, *pro* and *pol* genes. A) *gag* and *pol* genes of Moloney murine leukaemia virus (MuLV) are in-frame and separated by a stop codon. B) Overlapping of the 5' end of the *gag* gene with the 3' end of the *pol* genes of Rous sarcoma virus (RSV) and human immunodeficiency virus type-1 (HIV-1). C) Separation of the *gag* and *pol* genes of human T-cell leukaemia virus type 1 (HTLV-1) and mouse mammary tumour mouse (MMTV) by a overlapping *pro* gene.

Jacks *et al* (1988) have proposed a model for ribosomal frameshifting which consists of a heptanucleotide 'slippery' sequence with a stem-loop structure located 3' of this sequence. According to the proposed model, simultaneous slippage of the ribosomes in the A and P positions occurs in the -1 direction during translation. Mutational analysis of the suspected slippery sequence of Rous sarcoma virus (RSV) demonstrated that the heptanucleotide sequence, AAAUUUA is involved in ribosomal frameshifting (Jacks *et al*, 1988). The *gag* and *pro* genes were expressed *in vitro* in rabbit reticulocyte lysate following *in vitro* RNA synthesis. SDS-PAGE analysis

revealed the presence of a *gag* encoded polypeptide and a *gag-pol* fusion polypeptide, indicating the occurrence of ribosomal frameshifting. Jacks *et al* (1988) suggested that a heptanucleotide sequence consisting of two homopolymeric triplets is the site at which ribosomal frameshifting occurs, following comparison of presumed frameshift sequences of a number of retroviruses. Mutational analysis of the slippery sequence allowed the identification of the nucleotides important in the frameshifting event. In the heptanucleotide sequence $_{-4}$AAAUUUA$_{+3}$ (if numbering begins at -4 through to +3) mutation of the U nucleotides at positions -1, +1, and +2 to either A, C or G to reduced the efficiency of ribosomal frameshifting by greater than 10-fold, whereas mutation of the A nucleotides at positions -4, -3 and -2 reduced frameshifting 5-fold. Mutation of the A residue at position +3 to either C or G had no effect on frameshifting. However, changing this nucleotide to a U residue resulted in a 2-fold increase in frameshifting. Mutations which disrupted a stem-loop structure located 3' of this slippery sequence resulted in a 10-fold decrease in frameshifting. It is possible that the stem-loop structure could be involved in binding of a ribosomal protein, a soluble elongation factor or RNA which interferes with the stability of the ribosomal or tRNA component of the translational complex (Jacks *et al*, 1988). Furthermore sequences which are complementary to and located downstream of the stem-loop structure can interact to form a possible pseudoknot structure.

Ribosomal frameshifting mechanisms in which the retrovirus *pol* gene overlaps the *gag* gene have been described for human immunodeficiency virus type-1 (HIV-1) and feline immunodeficiency virus (FIV). Wilson *et al* (1988) identified the slippery region of HIV-1 as a 16 nucleotide overlap between the HIV-1 *gag* and *pol* genes in an *in vitro* translation system. The data indicated that the frameshifting event most probably involves a run of homopolymeric U residues. The construct containing the 16 nucleotide overlap region lacked a possible stem-loop structure that could have formed 3' of the slippery sequence if a longer overlapping region was used. Thus,

ribosomal frameshifting in HIV-1 appears to be independent of a stem-loop or pseudoknot structure. Reil *et al* (1993) also demonstrated a ribosomal frameshift event to occur between the *gag* and *pol* genes of HIV-1. Constructs were made in which the *gag* and *pol* genes were replaced by the β-galactosidase and luciferase encoding genes respectively. The β-galactosidase and luciferase genes were separated by a 59 base pair nucleotide sequence of HIV-1 that contained the slippery sequence (UUUUUUA) and a stem-loop structure located downstream of this sequence. Expression of the luciferase gene demonstrated that a frameshift event had occurred whereas mutation of the slippery sequence to CUUCCUC reduced the expression of luciferase by 100-fold, and deletion of the stem-loop structure resulted in only a 3-fold decrease in luciferase expression. These results indicate the slippery sequence plays a major role in the frameshifting event with the stem-loop structure acting as a positive modulator of ribosomal frameshifting (Reil *et al*, 1993). A complication with this hypothesis was the implication that mutagenesis of a heptanucleotide, GUUUUUA located within the coding region of the luciferase gene to the slippery sequence of HIV-1 (UUUUUUA) had occurred. Furthermore, the report implies that only the stem-loop structure of HIV-1 was inserted downstream of this sequence. Thus, the sequence located around this region the slippery sequence which is responsible for frameshifting may not represent true HIV-1 nucleotide sequences. Parkin *et al* (1992) transfected avian and Cos-7 cells with plasmid constructs containing the full-length HIV-1 *gag* and *pol* genes under the control of the cytomegalovirus enhancer-promoter with the HIV-1 *rev* response element located 3' of the *pol* gene. The *rev* response element was inserted to allow trans-activation of *gag* and *pol* by *rev*. The HIV-1 *rev* was supplied by co-transfection with a construct containing the *rev* gene also under the control of the cytomegalovirus enhancer-promoter. Transfection of cells with constructs in which the stem-loop structure sequence had been mutated or deleted showed a 4-7 fold decrease in frameshifting efficiency. *In vitro* expression, using rabbit reticulocyte lysate, of these

constructs in which expression was under the control of the SP6 promoter also showed that frameshifting was reduced by 2 to 3-fold (Parkin *et al*, 1992). This confirmed that the stem-loop structure acts as positive modulator of ribosomal frameshifting in HIV-1.

Using mutational analysis, Morikawa and Bishop (1992) identified the FIV slippery sequence to be near the heptanucleotide GGGAAAC. Recombinant baculovirus containing the FIV *gag* gene and the amino terminus of the protease protein of the *pol* gene was used to show frameshifting in baculovirus infected cells (Morikawa *et al*, 1991). In a construct in which a stop codon was inserted 5' of the slippery site , no product arising from a frameshifting event could be detected in insect cells infected with recombinant baculovirus. However, a protein arising from a frameshift event was observed when a stop codon was inserted 3' of the suspected slippery site. Furthermore, insertion of spacer nucleotides within the putative pseudoknot forming sequence located 3' of the slippery sequence had dramatic effects upon ribosomal frameshifting. These results suggest a role for pseudoknots in ribosomal frameshifting (Morikawa and Bishop, 1992).

Expression of the *pol* gene in retroviruses where the *gag* and *pol* genes are separated by an overlapping *pro* gene involves a double ribosomal frameshifting event. Jacks *et al* (1987) inserted the *gag-pro-pol* genes of mouse mammary tumour mouse (MMTV) in-frame with the 3' end of the *gag* gene of RSV which was under the control of the SP6 promoter. The RSV *gag* provided the initiation codon and also acted as an antigen for immunoprecipitation of the resultant translation products. Immunoprecipitation of *in vitro* synthesised polypeptides from this construct showed the presence of *gag-pro-pol* fusion polypeptide, indicating the occurrence of a double ribosomal frameshift event. Synthetic oligonucleotides representing overlapping regions containing the slippery sequences of the *gag/pro* or *pro/pol* genes were inserted between the RSV *gag* and HIV-1 *pol* genes. Expression of the *gag-pol* fusion

polypeptides was not observed, indicating the absence of either the sequence or structurally important determinants for frameshifting. Nam *et al* (1993) cloned the *gag-pro-pol* genes of human T-cell leukaemia virus type 1 (HTLV-1) under the control of the SP6 promoter. Only synthesis of polypeptides corresponding to expression of the *gag* gene and *gag-pro* fusion protein arising from ribosomal frameshifting between the *gag-pro* genes were observed *in vitro*. Insertion of a single nucleotide in the *gag-pro* gene overlap region resulted in the *gag* and *pro* genes becoming in-frame, but the *pol* gene remained in the -1 reading frame with respect to the *pro* gene. Expression of this mutant construct *in vitro* resulted in expression of the *gag/pro-pol* fusion product which was immunoprecipitated with rabbit antisera directed against the C terminus of the HTLV-1 integrase (IN) protein expressed from the *pol* gene. Mutation of the HTLV-1 heptanucleotide sequence UUUAAAC to UUUAAGC severely inhibited frameshifting indicating this to be the slippery sequence.

Wills *et al* (1991) described a read-through mechanism for Moloney murine leukaemia virus (MuLV) dependent upon a pseudoknot structure. In MuLV, although the *gag* and *pol* genes are in-frame, the ORF's are separated by a UAG stop codon and the *pol* ORF does not contain an AUG initiation codon. Approximately 5-10% of ribosomes have been estimated to read-through this stop codon although the normal rate for ribosomal read-through is less than 10% (Capone *et al*, 1986 referenced within Wills *et al*, 1991). Additionally, mutation of the UAG codon of MuLV to either UAA or UGA had no effect on the percentage of ribosomal read-through (Feng *et al*, 1989) although changing the UAG codon to CAG, encoding a glutamine residue inhibited viral release (Felsenstein and Goff, 1988). Cells carrying the mutant provirus constitutively expressed the *gag-pol* fusion polypeptide with expression of *gag* polypeptide completely inhibited. The *gag-pol* fusion polypeptide was not processed by the protease product of the *gag* gene. Complementation of the mutant provirus by transfection into cells expressing the MuLV protease resulted in a

low level of virus release. The authors suggest the high levels of the *gag-pol* fusion polypeptide may have inhibited virus release and that a correct ratio between the *gag* and *gag-pol* fusion polypeptides was required for efficient maturation of the virus (Felsenstein and Goff, 1988). Wills *et al* (1991) identified the presence of a possible pseudoknot structure 3' of this termination codon. The stop-codon and pseudoknot encoding regions of MuLV were fused 5' of the lacZ gene reporter gene. Although disruption of the stem structure of the pseudoknot by site-directed mutagenesis reduced the efficiency of the read-through mechanism, altering 4 G residues located 3' of the stem-loop structure, thought to pair with C residues located within the loop of the stem-loop structure of the pseudoknot, to C residues resulted in abolition of ribosomal read-through. Changing the C residues to T (U) residues within the loop resulted in decreased inhibition of the read-through mechanism. It was suggested that 4 T residues found within the lacZ gene could have base paired with these mutated nucleotides resulting in the formation of the pseudoknot (Wills *et al*, 1991). Thus, the pseudoknot appears to play a role in ribosomal read-through as well as in ribosomal frameshifting.

Ribosomal frameshifting is not restricted to the retroviruses. Tzeng *et al* (1992) have demonstrated a frameshift event to occur in the overlap region of the *cap* and *pol* genes of *Saccharomyces cerevisiae* (yeast) virus type 1 (ScV). The slippery sequence was identified as GGGUUUA with a possible pseudoknot structure located 3' of this sequence. Site-directed mutagenesis of the sequence involved in the pseudoknot resulted in a reduced efficiency for frameshifting. Expression of the *cap* gene and the *cap-pol* overlap region fused to the α-peptide of the β-galactosidase gene in either wheat germ extract, *E. coli* or yeast demonstrated a ribosomal frameshift event to occur. The efficiency of frameshifting was 3.5% in wheat germ extract and 20-29% in yeast cells. These results demonstrate the influence of the expression system used upon the frameshifting event. Using a similar expression system, in which

a 396 bp fragment containing the ScV slippery sequence and the pseudoknot structure were fused 5' of the β-galactosidase gene under the control of the SP6 promoter such that the *cap* reading frame is in frame with the AUG located within the multiple cloning site and the *pol* reading frame was in frame with the β-galactosidase gene, Tu *et al* (1992) identified the ribosomal pause site during the frameshifting event. The 'heel-printing' technique was adopted to identify these pause sites in *in vitro* transcribed/translated constructs. *In vitro* synthesised SP6 transcripts translated in the presence of rabbit reticulocyte lysate were treated with micrococcal nuclease. Only the regions of RNA which bound proteins or ribosomes were protected from digestion by steric hindrance. Purified, undigested RNA fragments were hybridised to negative sense ssDNA encoding the 396 bp fragment used above. 5' extension of a radiolabelled primer that hybridises upstream of the slippery site was achieved by T4 DNA polymerase. Since T4 DNA polymerase lack 5'→3' exonuclease activity, extension of the primer was strongly terminated when the T4 DNA polymerase encountered the RNA:DNA hybrid region. This allowed identification of the ribosomal pause site to be mapped just 5' of the slippery sequence. Similar analysis of mutant constructs containing disrupted pseudoknot structures showed a decrease in termination of primer extension in this region. The results indicate a slowing down or pausing of ribosomes during translation 5' of the pseudoknot structure (Tu *et al*, 1992).

Prufer *et al* (1992) demonstrated ribosomal frameshifting to occur in the single stranded, positive sense RNA plant virus, potato leafroll luteovirus. The overlapping regions of the ORF2a and ORF2b genes were fused internally to the *E. coli* β-glucuronidase (GUS) gene. Site-directed mutagenesis was employed to identify the slippery sequence as UUUAAAU. Mutation analysis of the slippery sequence had little effect on the efficiency of ribosomal frameshifting. Although, the sequence 3' of this slippery sequence is capable of forming a stem-loop structure, it is not capable of forming a pseudoknot. Ribosomal frameshift events have also been suggested in

expression of reverse transcriptase in retrotransposons (Mellor *et al*, 1985, Wilson *et al*, 1986, Clare *et al*, 1988) and has been described for *E. coli dnaX* (Flower *et al*, 1990, Tsuchihashi and Kornberg, 1990, Tsuchihashi, 1991)

Brierley *et al* (1987, 1989) described ribosomal frameshifting in the 1a and 1b (formerly called F1 and F2 respectively) ORF's of IBV. The overlapping frameshift region was cloned into a marker gene and expressed both *in vitro* and *in vivo*. Deletion analysis of the suspected slippery sequence identified the motif UUUAAAC to be important in frameshifting. Site-directed mutagenesis of the UUUA motif of the slippery sequence to CUUA reduced frameshifting by 10-fold. Mutagenesis of AAAC to ACAC reduced frameshifting dramatically and mutagenesis of AAAC to AUAC completely abolished frameshifting. Furthermore, mutagenesis to destabilise a stem-loop structure with the potential to form a pseudoknot reduced frameshifting efficiency whereas mutagenesis of nucleotide sequences downstream of the stem-loop structure thought to base pair with the loop of the stem-loop structure reduced frameshifting 10-fold. These results suggest a crucial role for the pseudoknot in ribosomal frameshifting (Brierley *et al*, 1989). Further investigation of the role of the IBV pseudoknot in ribosomal frameshifting, by site-directed mutagenesis of the IBV pseudoknot determined that so long as the overall structure of the pseudoknot was maintained, alteration of the nucleotide sequence within the pseudoknot had little effect of frameshifting. Replacement of the pseudoknot with a simple stem-loop structure of similar stem length and composition resulted in a marked reduction of frameshift efficiency (Brierley *et al*, 1991). Further mutagenesis of the IBV slippery sequence upon ribosomal frameshifting determined that in the heptanucleotide sequence XXXYYYN corresponding to the IBV slippery sequence, UUUAAAC, triplets of A, C, U or G were functional in the ribosomal P-site (X residues) but that G and C residues were non-functional in the ribosomal A-site (Y residues). They

determined the identity of the nucleotide at position N to be crucial with the order of preference being C > A ≅ U>>G (Brierley *et al*, 1992).

In summary, the general requirements for efficient ribosomal frameshifting involves the presence of overlapping reading frames, a slippery sequence corresponding to the heptanucleotide sequence XXXYYYN and presence of a stem-loop or pseudoknot structure 3' of this slippery sequence that may function as a enhancer of the frameshifting event. Although, in the examples presented above frameshifting occurs predominantly in the -1 direction, frameshifting into the +1 frame has been observed when the heptanucleotide sequence consisted of a run of homopolymeric residues following mutagenesis of the slippery sequences of IBV (Brierley *et al*, 1992). Mutagenesis of the slippery sequence of IBV to U_6N or U_8 resulted in increased frameshifting in the +1 or -2 reading frames. Disruption of the pseudoknot structure downstream of these (U_6N or U_8) slippery sequences gave different results. Whilst frameshifting in the construct containing the U_6N sequence was reduced, no reduction in frameshifting efficiency was observed when the pseudoknot was disrupted in the construct containing the U_8 slippery sequence. Brierley *et al* (1992) suggest the pseudoknot structure to influence the direction of the frameshift event.

In relation to PVM strain 15 G gene, the reading frame initiating at AUG_{83} overlaps the reading frame of AUG_{183}. A change of reading frame in the -1 direction would allow access to the reading frame of AUG_{183}, resulting in the addition of a cytoplasmic tail to the G protein. Sequence analysis of the 5' end of the PVM strain 15 G gene reveals the presence of a termination codon, UAA, located in-frame but upstream of UAG_{183} at position 168. Thus, any frameshift event that could occur must be between this termination codon and that for the small ORF of the gene at position 188. Although the G residue in position N of the heptanucleotide XXXYYYN was found to considerably inhibit frameshifting in IBV (Brierley *et al*, 1992), a G

residue in this position following mutagenesis of the slippery sequence of RSV (Jacks *et al*, 1988) had little effect upon frameshifting efficiency. Furthermore, the slippage sequence for the *gag-pro* genes of MMTV is GGAUUUA (Moore *et al*, 1987, Jacks *et al*, 1987) which does not conform to the heptanucleotide sequence suggested by Jacks *et al* (1988) and Brierley *et al* (1992). Other examples of slippery sequences which do not conform to the suggested heptanucleotide sequence are GGAUUUU and GUUAAAC corresponding to the p27-*pol* and 1a-1b overlap regions of red clover necrotic mosaic virus RNA-1 (Xiong and Lommel, 1989) and equine arteritis virus (den Boon *et al*, 1991) respectively. Thus, it is possible that the heptanucleotide sequences UAAAAAG or AACCAAA located between the two termination codons (at nucleotide positions 168 and 186) in the PVM strain 15 G gene may represent possible slippery sites. No pseudoknot structure is located 3' of this suspected slippery sequence for PVM although a stem-loop structure is present immediately after the second suspected slippery site (AACCAAA; Fig. 7.2a). However, a pseudoknot structure is possible in which the stem-loop structure is located immediately preceding the UAAAAAG sequence of the G gene (Fig. 7.2b). Brierley *et al* (1989) determined the length of the 6 nucleotide long sequence (spacer sequence) of IBV located between the slippery sequence and stem-loop structure to be essential for efficient frameshifting. Similarly, Morikawa and Bishop (1992) determined the requirement for the spacer sequence of FIV to be precisely 8 nucleotides. Addition or deletion of codon triplets in IBV or FIV resulted in a decreased efficiency of ribosomal frameshifting. However spacer sequences of 1, 4, 4, 5, 6, and 6 nucleotides have been identified for RSV (Jacks *et al*, 1988) yeast double-stranded RNA virus (Tzeng *et al*, 1992), *pro-pol* overlap region of MMTV (Jacks *et al*, 1987), potato leafroll luteovirus (Prufer *et al*, 1992), *gag-pro* genes of MMTV (Jacks *et al*, 1987) and HTLV-1 (Nam *et al*, 1993) respectively. Even within the same virus, MMTV, the length of the spacer sequence is not absolute between overlapping regions of the *gag-pro* and *pro-pol*

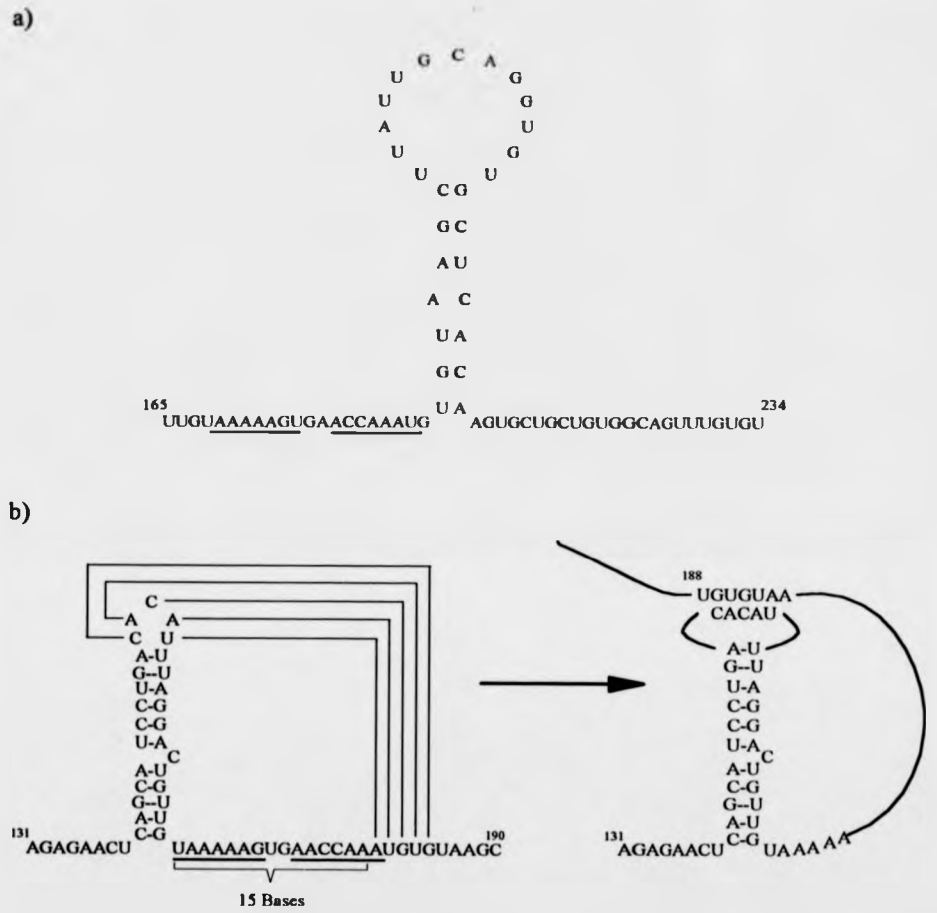


Fig. 7.2 : Secondary structure formation within PVM strain 15 G gene.

Diagrammatic representation of a) possible stem-loop structure located 3' of the suspected slippery sequences (underlined) and b) pseudoknot structure located within the 5' end of the G gene of PVM strain 15 that encompasses the possible slippery sequence (underlined). The loop of the stem-loop structure is capable of pairing with a sequence located 3' of the slippery sequence. Numbers shown against individual nucleotides correspond to nucleotide positions shown in Fig. 4.4.

genes. Furthermore, HIV-1 frameshifting has been reported to occur in the absence of any such structures 3' of the slippery sequence (Wilson *et al*, 1988). It is not known whether a pseudoknot structure 5' of slippery sequence enhances frameshifting. Thus , it may be possible that slippery sequence located within a possible pseudoknot may function correctly in a frameshifting event.

Thus, an investigation into the possibility of a ribosomal frameshifting event occurring at the 5' end of PVM strain 15 G gene was undertaken. The approach adopted involved mutagenesis of the AUG codons located at positions 83, 183 and 246 of the G gene of PVM strain 15 and fusion of these constructs 5' of the CAT reporter gene. Expression of the CAT gene was then assayed in transfected cells. Presence of CAT activity found in cells transfected with a construct in which the AUG at position 183 and not 83 has been changed would indicate presence of a ribosomal frameshifting event occurring, since initiation from AUG_{83} could not result in expression of the CAT reporter gene by any other process.

7.2. RESULTS :

7.2.1. Purification of PVM strain 15 G gene 5 end :

Clone G(3) containing the full-length G gene of PVM strain 15 was digested with restriction endonuclease *PvuII* to release a 600 bp fragment containing the 5' end of the G gene. *PvuII* cuts at positions 563 and 976 of the nucleotide sequence of PVM strain 15 G gene and also 5' of the binding site for the universal (-40) primer and 3' of the multiple cloning site of pUC13 plasmid vector (Fig. 7.3). (Digestion of pUC13 vector with *PvuII* alone would result in the release of a fragment of approximately 300 base pairs.) These DNA fragments was then used in subsequent mutagenesis of the AUG codons of the 5' G gene end located within this fragment.

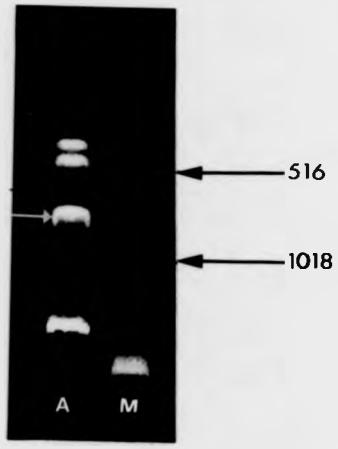


Fig. 7.3 : *PvuII* digested clone G(3).

1% LMP agarose gel stained with ethidium bromide showing the release of the 600 bp DNA fragment containing the 5 end of PVM strain 15 G gene (arrowed) following digestion with restriction endonuclease *PvuII*. M represents the DNA marker lane.

7.2.2. Mutagenesis of PVM strain 15 G gene AUG codons :

Digestion of clone G(3) plasmid DNA with *PvuII* resulted in the release of a fragment containing the 5' end of PVM strain 15 G gene containing AUG codons at positions 83, 183 and 246. Since ribosomal frameshifting events were to be investigated in the region lying between AUG_{83} and AUG_{183}, AUG_{246} was altered to UUG_{246} resulting in change of a Met to Leu to prevent possible in-frame internal initiation at this point. Mutagenesis of the specific AUG initiation codon was achieved

using PCR with synthetic oligonucleotides which contained a mutation at the required point. Amplification of PCR products was performed in the presence of Vent™ DNA polymerase. Table 7.1 shows the nucleotide sequences of the oligonucleotides and primers used to mutate the AUG codons of the 5' end of PVM strain 15 G gene.

7.2.2.1. MUTAGENESIS OF AUG_{183}:

Fig. 7.4 shows the PCR product (-40/G5) following amplification of the first 209 nucleotides which resulted in alteration of AUG_{183} to UUG_{183} using the

Oligonucleotides	Nucleotide sequence	Position	Sense	Restriction sites
Universal primer (-40)	GTTTTCCCAGTCACGAC		+	None
G5	GAGCACACCTGCAATAAG CTTACACAATTGGTTCAC	209-174	-	*HindIII*
G6	CCGGATCCCGGGAAAACA GAATACAATATGACCCCC AC	259-234	-	*HaeIII, BamHI, SmaI*
G7	GTTCCTTCCAATACTAGG TGAA	94-73	-	None
GRev7	TTCACCTAGTATTGGAAG GAACT	73-95	+	None

Table 7.1 : Primers and oligonucleotides used in mutagenesis of G gene AUG codons.

Nucleotide sequence of oligonucleotides and primers used in mutagenesis of AUG_{83}, AUG_{183} and AUG_{246} and amplification of the 5' end sequences of the PVM strain 15 G gene. The nucleotide coding for the mutational change is shown in bold and underlined whereas those not found within the G gene sequence are underlined.

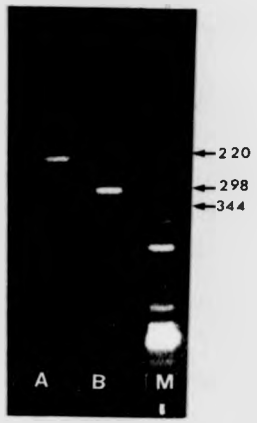


Fig. 7.4 : PCR product with altered AUG_{183}.

2% agarose gel stained with ethidium bromide showing PCR product (A) following amplification of *PvuII* digested PVM strain 15 G gene 5' end with the universal (-40) primer and oligonucleotide G5. M represents the DNA marker lane whilst lane B contains DNA fragment not relevant to this study.

universal primer and oligo G5. The PCR product was digested with restriction endonuclease *XbaI* (site located 5' of the G gene start) and *HindIII* (site located within oligo G5), purified from a 1% low melting point agarose gel and cloned into *XbaI* and *HindIII* digested pBS. Inserts from 5 individual colonies were released following analysis by the miniprep DNA protocol, digested with *XbaI* and *HindIII*, sub-cloned into M13 mp18 and the nucleotide sequence determined (data not shown).

7.2.2.2. MUTAGENESIS OF AUG_{83} :

Fig. 7.5 shows the procedure adopted for changing initiation AUG_{83} to ATT_{83} which involved the use of back-to-back oligo's carrying the mutated nucleotide

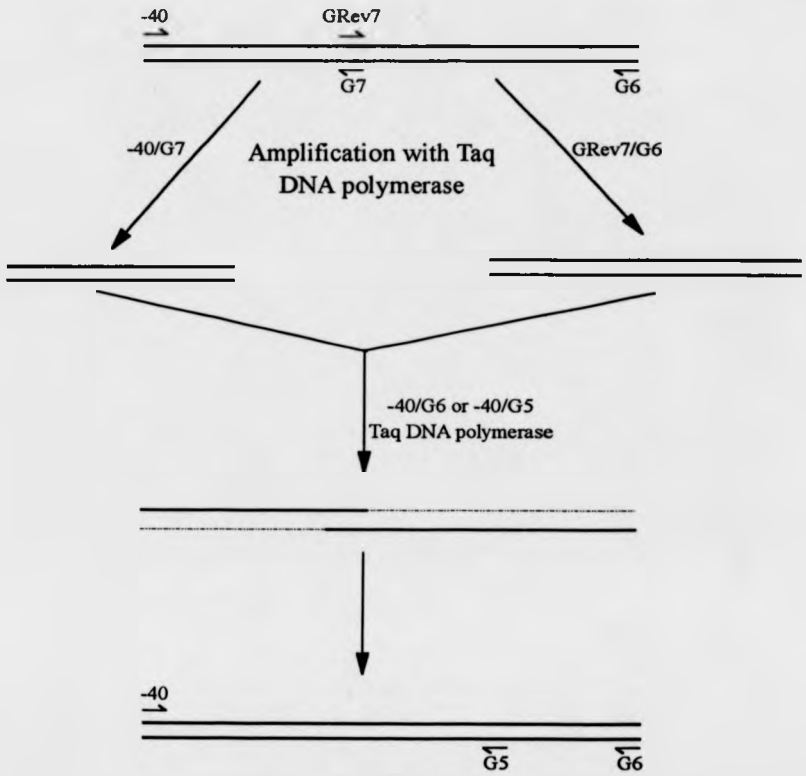


Fig. 7.5 : Diagrammatic representation of protocol used for alteration of AUG_{83}, AUG_{183} and AUG_{246}.

Diagrammatic representation of the two-step PCR process involved in mutagenesis of AUG_{83} and AUG_{246} within the same construct. The first step involves amplification of two segments of the 5' end of PVM strain 15 G gene with the -40 primer/oligo G7 and oligo's Grev7/G6. The second step involves amplification of a mixture of the two PCR products (obtained from step one) with a combination of the -40 primer and oligonucleotides G6 and G5.

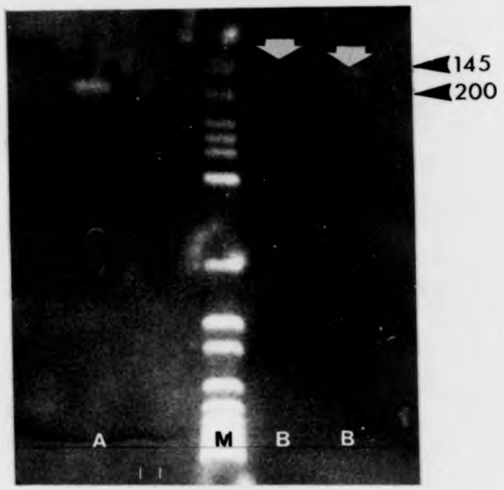


Fig. 7.6 : PCR products containing altered AUG_{83} and AUG_{246}.

1% TAE agarose gel stained with ethidium bromide showing purification of PCR products amplified using oligonucleotides a) GRev7 and G6 and b) GR7 and -40 primer (arrowed). M represents the DNA marker lane.

that hybridised to AUG_{83}. Two separate PCR reactions were carried out on dilutions of the 5' G gene end purified following *PvuII* digestion. The first reaction involves amplification of the first 94 nucleotides of PVM strain 15 G gene with the universal primer and oligo G7. Amplification with these primers results in alteration of AUG_{83} to AUU_{83}. A second PCR reaction using oligonucleotides GRev7 and G6 resulted in amplification of nucleotides 73 to 259 in which AUG_{83} and AUG_{246} had been changed to AUU_{83} and UUG_{246} respectively.

The PCR products obtained following amplification with the universal primer/G7 and Grev7/G6 were purified from 1% TAE agarose gel (Fig. 7.6) using the GeneClean™ kit (Bio 101, La Jolla, California, USA.). A second PCR reaction of

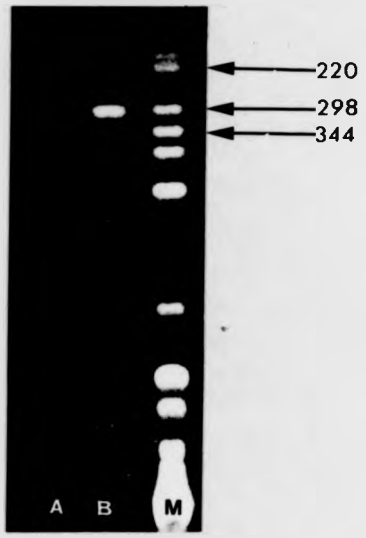


Fig. 7.7 : PCR products containing altered AUG_{83}, AUG_{183} and AUG_{246}.

1% LMP agarose gel stained with ethidium bromide showing purification of (A) *XbaI* and *HindIII* digested G7/G5 (see text) and (B) *XbaI* and *SmaI* digested G7/G6 PCR products. M represents the DNA marker lane.

diluted aliquots of the above PCR products were carried out and the full-length PCR product amplified using the universal primer and oligo G6. The PCR product (G7/G6, Fig. 7.7) obtained from this reaction was digested with restriction endonucleases *XbaI* and *SmaI*, purified from a 1% low melting point agarose gel and cloned into *XbaI* and *SmaI* cut pBS. This contains the first 259 nucleotides of PVM strain 15 G gene in which AUG_{83} and AUG_{246} had been changed to ATT_{83} and TTG_{246} respectively. The *XbaI* restriction site was present 5' of the G gene start in clone G(3) whereas *SmaI* restriction site was located within oligo G6. Colonies containing these constructs were screened by digestion with *EcoRI* and *PstI* of DNA prepared by the miniprep method.

One of the released inserts was subcloned into *EcoRI* and *PstI* digested M13 mp19 and the nucleotide sequence determined (data not shown). Unfortunately, it appeared that pBS digested with *SmaI* lost 3 nucleotides, possibly due to contamination of during digestion pBS vector (data not shown).

7.2.2.3. MUTAGENESIS OF BOTH AUG_{83} AND AUG_{183} :

Mutations of AUG_{83} and AUG_{183} were introduced within the same DNA fragment by amplification with the universal primer and oligonucleotide G5 during the second step of the mutagenesis procedure described above. Fig. 7.7 shows the PCR product (G7/G5) purified from a 1% low melting point agarose gel following digestion with restriction endonucleases *XbaI* and *HindIII*. The digested PCR product was cloned into *XbaI* and *HindIII* digested pBS and subsequently subcloned into M13 mp18 for nucleotide sequencing. Fig. 7.8 shows the nucleotide sequence of two constructs in which AUG_{83} and AUG_{183} have been changed to AUU_{83} and UUG_{183} respectively.

7.2.3. Purification of the β-galactosidase gene :

Oligonucleotide G6 was designed to enable the insertion of the CAT gene lacking its initiation codon to be inserted in frame with AUG_{183}, following digestion of plasmid pSV2-CAT with restriction endonuclease *MnlI*, 3' of the *SmaI* site located within oligonucleotide G6. According to the manufacturers of *MnlI*, this enzyme cut at the site $CCTC(N)_7$ leaving a blunt end fragment. However, since the initial design of oligonucleotide G6 the restriction site for *MnlI* has been modified to :

$$\begin{array}{l} 5'CCTC(N)_{7}\downarrow \\ 3'GGAG(N)_{6}\downarrow \end{array}$$

which leaves a single base 3' overhang. Thus, blunting of this *MnlI* cut CAT fragment could only be achieved by removal of this one base overhang which would result in the open reading frame of *MnlI* digested CAT gene to be out of frame

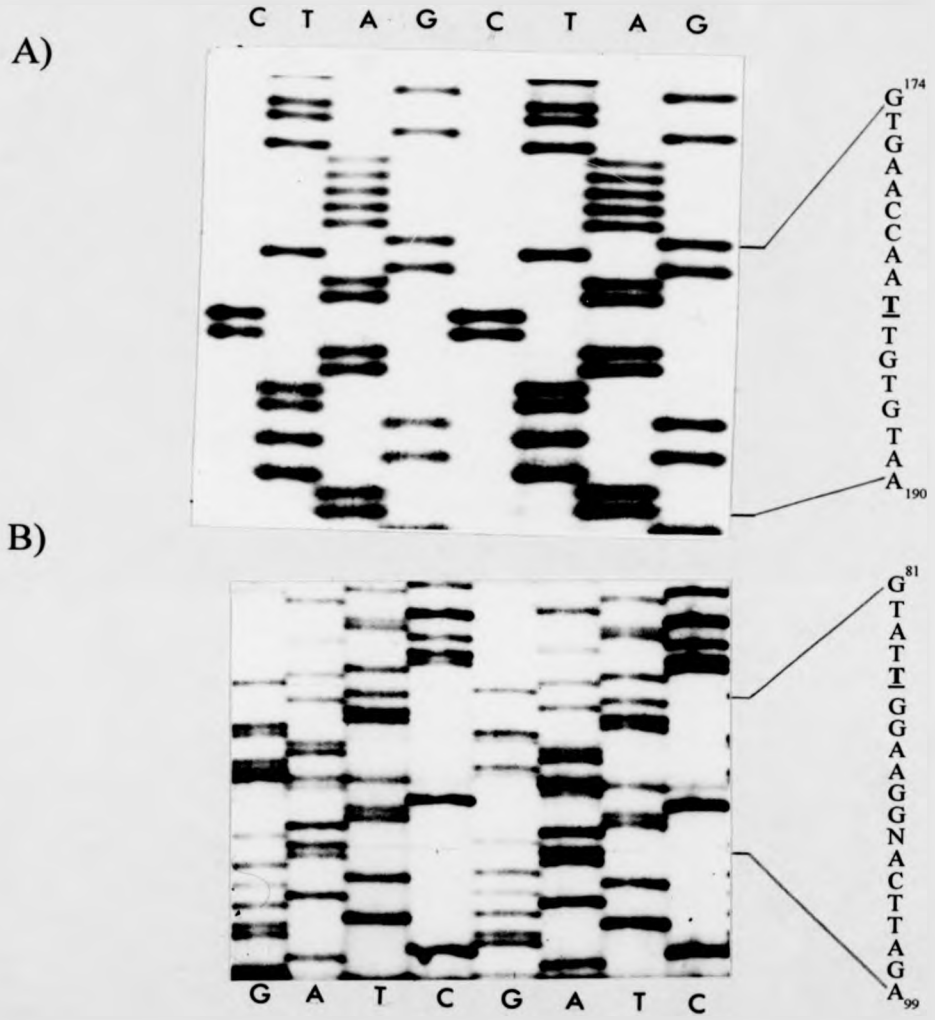


Fig. 7.8 : Nucleotide sequence of PCR product containing altered AUG_{83} and AUG_{183}.

Autoradiograph showing the nucleotide sequence of a clone of the G7/G5 PCR product in which initiation codons at positions A) 183 and B) 83 of the G gene of PVM strain 15 have been altered. Mutated nucleotides are underlined and shown in bold. Numbers shown against individual nucleotides correspond to nucleotide positions shown in Fig. 4.4.

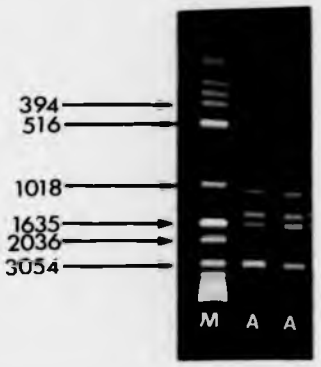


Fig. 7.9 : Purification of *E. coli* β-galactosidase gene.

SalI, *XmnI* and *SmaI* Restriction endonuclease digest of plasmid pMC1871 (A) from which the *E. coli* β-galactosidase gene (arrowed) was excised. M represents the DNA marker lane.

with AUG_{183}. Thus, it was decided to place the *E. coli* β-galactosidase gene, that lacked its initiation codon, downstream of AUG_{183}. Plasmid pMC1871 contains the *E. coli* β-galactosidase gene that lacks its initiation codon and allows insertion of DNA fragments at a *SmaI* site located at the 5' end of the gene. Unfortunately, ligation of *SmaI* cut pMC1871 and the DNA fragment amplified using oligonucleotide G6 does not place the β-galactosidase gene in frame with AUG_{183}. However, cloning of the β-galactosidase gene in frame with AUG_{183} can be achieved by excision of the gene with digestion with restriction endonucleases *SmaI* and *SalI* and cloned into a suitable restriction site located 3' of the cloned DNA fragment amplified using oligonucleotide G6. Fig. 7.9 show restriction endonuclease digest of pMC1871 with restriction endonucleases *SmaI*, *XmnI* and *SalI*. Plasmid pMC1871 was digested with *XmnI* so as to distinguish between the released β-galactosidase gene and the remainder of the

plasmid DNA since digestion with just *SmaI* and *SalI* would result in generation of fragments of approximately equivalent sizes. The released β-galactosidase gene was purified from a 1% low melting point agarose gel and has yet to be sub-cloned 3' of the cloned DNA fragment amplified using oligonucleotide G6.

7.3. DISCUSSION :

To investigate the possibility of the addition of a cytoplasmic tail to the N terminus of the G protein of PVM strain 15 by ribosomal frameshifting, AUG codons at positions 83, 183 and 246 of the PVM strain 15 G gene were mutated to AUU_{83}, UUG_{183} and UUG_{246} respectively. The use of a different oligonucleotides in the polymerase chain reaction was employed to amplify the 5' end of the G gene in which a combination of the AUG codons of the G gene were mutated. The mutation of AUG_{83} in a two step process was performed so that a PCR product representing the exact 5' end of the G gene only contained this mutated initiation codon. Mutation of the AUG_{246} was necessary so that expression of the β-galactosidase reporter gene was not possible from an initiation codon located downstream of the suspected slippery sequences in the G-β-galactosidase fusion protein. Furthermore, it was desirable to include the sequence located 3' of the putative slippery sequences within the fusion construct so that a potential stem-loop structure, which is known to act as a positive modulator of ribosomal frameshifting, could be formed.

Oligonucleotide G6 contains the base change that results in the mutation of AUG_{246} to UUG_{246}. This oligo also contains at its 3' end, restriction endonuclease sites for *SmaI*, *BamHI* and *HaeIII*. The sites were arranged so that insertion of a CAT reporter gene lacking its initiation codon could be fused into the *SmaI* site of this construct. This would have resulted in the CAT gene being in-frame with AUG_{183} and that expression of the CAT reporter gene could only occur from AUG_{183}. It was anticipated that since the CAT reporter gene contains no internal

PCR Product	AUG_{83}	AUG_{183}	AUG_{246}
-40/G5	✓	✗	✗
G7/G6	✗	✓	✗
G7/G5	✗	✗	✗
Clone G(3)	✓	✓	✗

Table 7.2 : Summary of mutated AUG codons within each G-CAT construct.

Table showing summary of oligonucleotides used in construction of fragments carrying a combination of mutated initiation codons.

HindIII restriction sites that the constructs containing the mutated AUG_{83} and/or AUG_{183} could be interchanged within the G-CAT fusion gene by digestion with *HindIII*, since this site is located 3' of AUG_{183}. However, a revision of the restriction site for endonuclease *MnlI* after the synthesis of oligonucleotide G6 prevented the successful fusion of the CAT gene that would have resulted in it being in frame with AUG_{183}. However, it is possible to fuse the β-galactosidase gene lacking its initiation codon 3' of the G7/G6 construct. Expression of the β-galactosidase gene, lacking its first initiation codon, from internal initiations does not result in expression of a functional protein (D. Hodgson, personal communication). However, one drawback is that fusion of the β-galactosidase gene into the *SmaI* site of oligo G6 does not leave it in-frame with AUG_{183}. Manipulation of the plasmid vector specific restriction sites is necessary to circumvent this problem. Table 7.2 shows the various constructs containing mutated AUG codons that would have been fused to the reporter gene.

In summary. expression of the reporter gene following fusion 3' and in frame of AUG_{183} of the G7/G6 construct would only occur if AUG_{183} was present whereas in the absence of AUG_{183}, no expression of the reporter gene should be detected. However, in the construct which contains AUG_{83} but not AUG_{183}, if reporter gene expression is detected in transfected cells, and since AUG_{83} is out-of-frame with AUG_{183}, expression could only occur by a phenomenon such as ribosomal frameshifting. No expression should be detected from the construct in which both AUG_{83} and AUG_{183} have been mutated. The mutagenesis of the different AUG codons of PVM strain 15 G gene represents the first step towards investigating ribosomal frameshifting within a negative strand virus.

CHAPTER 8

GENERAL DISCUSSION

8.1. DISCUSSION :

To characterise the molecular basis of the different pathogenicities of two strains of PVM molecular analysis of the genes encoding the two large surface glycoproteins was performed. PVM Strain 15, the non-pathogenic variant, has been passaged, to our knowledge, entirely in tissue culture whereas PVM strain J3666, the pathogenic variant has been passaged entirely in mice. Growth of the two strains in tissue culture is extremely slow with PVM strain 15 showing cytopathic effect much earlier than PVM strain J3666 when cells are infected *in vitro* at the same multiplicity of infection and strain J3666 is a naturally temperature sensitive virus with no growth at 37°C (A. Easton, personal communication).

A comparison of the fusion protein gene sequences was considered to be essential since the pathogenicity of NDV isolates were attributed to amino acid changes within their fusion proteins (Glickman *et al*, 1988, Toyoda *et al*, 1987, Morrison *et al*, 1993). Nucleotide sequence analysis of NDV isolates characterised these changes as occurring in the vicinity of the F protein cleavage site. The cleavage of F protein into disulphide-bonded F_1 and F_2 subunits is crucial to the functionality of the protein. Thus, changes that alter the susceptibility of the cleavage site to cleavage by host proteases play an important role in determining the pathogenicity of the virus. Similarly, mutation of Sendai virus F protein surrounding the cleavage site alters the susceptibility of the protein to cleavage by extracellular proteases. For example, changing a single Arg residue to Ile at position 116 within the F_0 polypeptide of Sendai virus results resistance to cleavage by trypsin but confers susceptibility to cleavage by chymotrypsin (Blumberg *et al*, 1985). Similarly, mutation of amino acids around the F protein cleavage site of Sendai virus results in altered susceptibility to extracellular proteases (Itoh *et al*, 1987, Itoh and Homma, 1988, Ogasawara *et al*, 1992). Thus, nucleotide sequence analysis of the fusion glycoprotein of PVM strain J3666 was undertaken so that comparison with the nucleotide sequence of PVM strain

15 fusion protein might reveal the molecular basis for the pathogenesis of the two viruses (Chambers *et al*, 1992).

Of the 1657 nucleotides encoding the F protein of PVM strain 15 (Chambers *et al*, 1992), only six nucleotide changes were found. Of these, two did not result in alteration of the encoded amino acids whereas the remaining four resulted in amino acids changes. Furthermore, these changes were distributed throughout the F protein sequence. Although the effects of these changes were not characterised due to lack of appropriate MAb's it seems unlikely that they contribute significantly to the different pathogenicities of the two strains of PVM and may be a result of immune pressure since PVM strain J3666 has been passaged entirely in mice. Thus, these mutations may identify regions involved in eliciting an antibody response. However, the possibility can not be ruled out that one or more of these changes may be located within the 3-dimensional structure that forms the F_0 cleavage site, thus, influencing the susceptibility of the cleavage site of the F_0 polypeptide.

The second large glycoprotein of PVM is the G proteins which, by analogy to HRSV, acts as the virus attachment protein (Levine *et al*, 1987). The G protein of HRSV has been shown to be variable between isolates of the two subgroups and also isolates from within the same subgroup (Mufson *et al*, 1985, Anderson *et al*, 1985, Johnson *et al*, 1987b, Cane *et al*, 1991, Sullender *et al*, 1991). Thus it is possible that the attachment protein could play a role in the pathogenicity of both strains of PVM. A comparison of the nucleotide sequence of the G genes of both strains of PVM revealed several differences. Within the extracellular domain of the G protein only two nucleotide and resulting amino acid residue changes were observed. It is not known whether these result in some advantage for strain J3666 which could explain the enhanced pathogenesis. However, the most single important difference identified between the G proteins of PVM was the absence of amino acid sequences encoded by the major ORF at the N-terminus of the G protein of PVM strain 15 which

are present in the G protein of PVM strain J3666. These amino acid sequences, in PVM strain J3666, which would form the cytoplasmic tail of the G protein were found to be absent in the G protein of PVM strain 15. Thus, the G protein of strain 15 appears to lack a cytoplasmic tail that is present in strain J3666. The importance of this is unclear since it is generally believed that the cytoplasmic tail of the glycoproteins interact with the M proteins during virus maturation and its absence may detrimentally affect virus production.

Additionally, another difference highlighted by nucleotide sequence analysis of the G genes of PVM is the location of a short ORF located upstream of the major ORF within each virus. The major ORF of the PVM strain J3666 is encoded by an initiation codon located at position 83 of the nucleotide sequence with the short ORF initiating from an AUG located at position 29. The termination codon for this short ORF is located at position 65 of the nucleotide sequence for the G gene PVM strain 15. The insertion of an A residue within a stretch of A residues located between AUG_{83} and AUG_{182} (of strain J3666) results in the second initiation codon becoming out of frame in the G gene of PVM strain 15. Thus, the major ORF of the G gene of strain 15 is initiated at the AUG that is located at position 183 of the nucleotide sequence of PVM strain 15 G gene with a short ORF initiated by AUG_{83} which terminates 2 nucleotide bases downstream of AUG_{183}. However, the initiation codon of the short ORF located within the G gene of strain J3666 has been altered in strain 15 so that no ORF exists. Thus, the G genes of either strain of PVM possess two ORF's in which the second ORF codes for the major polypeptide even though they are initiated from separate AUG codons. A diagrammatic representation of the gene structure is shown in Fig. 4.9.

The presence of an upstream ORF in both G genes may play a role in down-regulation of G gene expression *in vivo*. *In vitro* translation of the G genes result in expression of polypeptides with sizes that are approximately 3K larger than

their predicted molecular weights. Expression of the G gene of PVM strain 15 from either a full-length clone or from a construct in which the first ORF has been deleted results in expression of polypeptides of similar size. However, although equivalent amounts of *in vitro* transcribed RNA, estimated from agarose gel electrophoresis, were used in the *in vitro* translation system, expression of the major polypeptide in the construct lacking the first ORF appeared to be increased when compared to expression from the clone containing the full-length G gene of strain 15, suggesting the presence of the first ORF to down-regulate expression from the major ORF. Furthermore, *in vitro* expression of the G gene of PVM strain J3666 results in synthesis of a polypeptide of approximately 47K and also a polypeptide of approximately 43K. This latter polypeptide is of similar size to that synthesised from constructs containing either the full-length or just the major ORF of PVM strain 15 G gene. Thus, this polypeptide may be a result of internal initiation within the G gene of strain J3666 and by analogy to the G protein of strain 15 may lack the cytoplasmic tail. It is not known whether this polypeptide is synthesised *in vivo* in PVM strain J3666 infected cells or whether it plays a role in virus maturation.

Expression of the G genes of PVM strains 15 and J3666 and the construct containing strain 15 G gene but lacking the first ORF were expressed in BS-C-1 cells using recombinant vaccinia virus expressing the T7 RNA polymerase. Immunofluorescence straining of transfected/infected cells revealed the surface location of the polypeptides synthesised from both full-length G genes and also from the construct containing the major ORF of the G gene of PVM strain 15. The results indicate that the G protein of strain 15 is processed and transported correctly within the cell even though it appears to lack a cytoplasmic tail. Thus, signals for the correct processing are not located within the cytoplasmic domain of the G protein J3666.

8.2. FUTURE WORK :

The identification of two translational products from *in vitro* expression of the G gene of PVM strain J3666 still leaves open the question as to whether the polypeptide synthesised due to internal initiation at AUG_{182} *in vitro* is also present in PVM strain J3666 infected cells.

Addition of a cytoplasmic tail to the N-terminus of the G protein synthesised from PVM strain 15 could be achieved by modification during translational of the mRNA. Sequences similar to those involved in ribosomal frameshifting which result in changing of the reading frame to that of -1 direction have been identified with relevant potential secondary structure elements in the RNA in the close proximity. The first steps in determining whether such a mechanism as ribosomal frameshifting occurs with the G gene of PVM strain 15 have been undertaken. Mutagenesis of the initiation codons coding for the major ORF or the short ORF of PVM strain 15 with a view to fusing these mutated sequences has been achieved with only ligation of these mutated sequences 5' of the *E. coli* β-galactosidase gene remaining. Subsequent, expression of these fusion constructs in tissue culture cells would help determine whether the frameshifting event occurs. Additionally, if ribosomal frameshifting was found to occur, then mutagenesis of the nucleotide sequence located within the two ORF's of PVM strain 15 would help identify the sequences responsible for such an event.

BIBLIOGRAPHY

Akerlind, B., Nörrby, E., Orvell, C. and Mufson, M. A. (1988). Respiratory syncytial virus : heterogeneity of subgroup B strains. *Journal of General Virology* **69**, 2145-2154.

Alkhatib, G. and Breidis, D. J. (1986). The predicted primary structure of the measles virus hemagglutinin. *Virology* **150**, 479-490.

Alkhatib, G., Massie, B. and Briedis, D. J. (1988). Expression of bicistronic measles virus P/C mRNA by using hybrid adenoviruses : Levels of C protein synthesized in vivo are unaffected by the presence or absence of the upstream P initiator codon. *Journal of Virology* **62**, 4059-4069.

Alkhatib, G., Richardson, C. and Shen, S. H. (1990). Intracellular processing, glycosylation, and cell-surface expression of the measles virus fusion protein (F) encoded by a recombinant adenovirus. *Virology* **175**, 2662-2670.

Anderson, K., Stott, E. J. and Wertz, G. W. (1992). Intracellular processing of the human respiratory syncytial virus fusion glycoprotein: Amino acid substitutions affecting folding, transport and cleavage. *Journal of General Virology* **73**, 1177-1188.

Anderson, L. J., Heirholzer, J. C., Tsou, C., Hendry, R. M., Fernie, B. F., Stone, Y. and McIntosh, K. (1985). Antigenic characterization of respiratory syncytial virus strains with monoclonal antibodies. *Journal of Infectious Diseases* **151**, 626-633.

Anderson, L. J., Bingham, P. and Hierholzer, J. C. (1988). Neutralization of respiratory syncytial virus by individual and mixture of F and G protein monoclonal antibodies. *Journal of Virology* **62**, 4232-4238.

Arbiza, J., Taylor, G., Lopez, J. A., Furze, J., Wyld, S., Whyte, P., Stott, E. J., Wertz, G., Sullender, W., et al. (1992). Characterization of two antigenic sites recognized by neutralizing monoclonal antibodies directed against the fusion

glycoprotein of human respiratory syncytial virus. *Journal of General Virology* **73**, 2225-2234.

Arumugham, R. G., Hildreth, S. W. and Paradiso, P. R. (1989). Fatty acid aclylation of the fusion glycoprotein of human respiratory syncytial virus. *Journal of Biological Chemistry* **264**, 10339-10342.

Bagai, S., Puri, A., Blumenthal, R. and Sarkar, D. P. (1993). Hemagglutinin-Neuramindase enhances F protein-mediated membrane fusion of reconstituted Sendai virus envelopes with cells. *Journal of Virology* **67**, 3312-3318.

Baker, S. C. and Moyer, S. A. (1988). Encapsidation of Sendai virus genome RNAs by purified NP protein during in vitro replication. *Journal of Virology* **62**, 834-838.

Baltimore, D. (1971). Expression of animal virus genomes. *Bacteriological Reviews* **35**, 235-241.

Banerjee, A. K. (1987). Transcription and replication of Rhabdoviruses. *Microbiological Reviews* **51**, 66-87.

Banerjee, A. K. and Barik, S. (1992). Minireview: Gene expression of vesicular stomatitis virus genome RNA. *Virology* **188**, 417-428.

Barik, S. and Banerjee, A. K. (1992). Sequential phosphorylation of the phosphoprotein of vesicular stomatitis virus by cellular and viral protein kinases is essential for transcription activation. *Journal of Virology* **66**, 1109-1118.

Barik, S. and Banerjee, A. K. (1991). Cloning and expression of the vesicular stomatitis virus phosphoprotein gene in *Escherichia coli*: Analysis of phosphorylation status versus transcriptional activity. *Journal of Virology* **65**, 1719-1726.

Barr, J., Chambers, P., Pringle, C. R. and Easton, A. J. (1991). Sequence of the major nucleocapsid protein gene of pneumonia virus of mice; sequence comparisons

suggest structural homology between nucleocapsid proteins of pneumoviruses, paramyxoviruses, rhabdoviruses and filoviruses. *Journal of General Virology* **72**, 667-685.

Baybutt, H. N. and Pringle, C. R. (1987). Molecular cloning and sequencing of the F and 22K membrane protein genes of the RSS-2 strain of respiratory syncytial virus. *Journal of General Virology* **68**, 2789-2796.

Bellini, W. J., Englund, G., Rozenblatt, S., Arnheiter, H. and Richardson, C. D. (1985). Measles virus P gene codes for two proteins. *Journal of Virology* **53**, 908-919.

Berthiaume, L., Joncas, J. and Pavilanis, V. (1974). Comparative structure, morphogenesis and biological characteristics of the respiratory syncytial (RS) virus and the pneumonia virus of mice (PVM). *Archive fur die gesamte virusforschung* **45**, 39-51.

Birnboim, H. C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* **7**, 1513-1523.

Black, B. L. and Lyles, D. S. (1992). Vesicular stomatitis virus matrix protein inhibits host cell-directed transcription of target genes in vivo. *Journal of Virology* **66**, 4058-4064.

Black, B. L., Rhodes, R. B., Mckenzie, M. and Lyles, D. S. (1993). The role of vesicular stomatitis virus matrix protein in inhibition of host-directed gene expression is genetically separable from its function in virus assembly. *Journal of Virology* **67**, 4814-4821.

Blondel, D., Harmison, G. G. and Schubert, M. (1990). Role of matrix protein in cytopathogenesis of vesicular stomatitis virus. *Journal of Virology* **64**, 1716-1725.

Blumberg, B. M., Giorgi, C., Rose, K. and Kolakofsky, D. (1985a). Sequence determination of the Sendai virus fusion protein gene. *Journal of General Virology* **66**, 317-331.

Blumberg, B. M., Giorgi, C., Roux, L., Rajin, R., Dowling, P., Chollet, A. and Kolakofsky, D. (1985b). Sequence determination of the Sendai virus HN gene and its comparison to the influenza virus glycoproteins. *Cell* **41**, 269-278.

Blumberg, B. M., Chan, J. and Udem, S. A. (1991) Function of 3' and 5' end sequences: In theory and practice in *The Paramyxoviruses*, ed. Kingsbury, D. W. (Plenum Press, New York and London), pp. 235-247.

Boeck, R., Curran, J., Matsuoka, Y., Compans, R. and Kolakofsky, D. (1992). The parainfluenza virus type 1 P/C gene uses a very efficient GUG codon to start its C' protein. *Journal of Virology* **66**, 1765-1768.

Bourgeois, C., Corvaisier, C., Bour, J. B., Kohli, E. and Pothier, P. (1991). Use of synthetic peptides to locate neutralizing antigenic domains on the fusion protein of respiratory syncytial virus. *Journal of General Virology* **72**, 1051-1058.

Brideau, R. J., Walters, R. R., Stier, M. A. and Wathen, M. W. (1989). Protection of cotton rats against human respiratory syncytial virus by vaccination with a novel chimaeric FG glycoprotein. *Journal of General Virology* **70**, 2637-2644.

Brierley, I., Bournsell, M. E., Binns, M. M., Bilimora, B., Blok, V. C., Brown, T. D. and Inglis, S. C. (1987). An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. *EMBO Journal* **6**, 3779-3785.

Brierley, I., Digard, P. and Inglis, S. C. (1989). Characterization of an efficient coronavirus ribosomal frameshifting signal: Requirement for an RNA pseudoknot. *Cell* **57**, 537-548.

Brierley, I., Rolley, N. J., Jenner, A. J. and Inglis, S. C. (1991). Mutational analysis of the RNA pseudoknot component of a coronavirus ribosomal frameshifting signal. *Journal of Molecular Biology* **220**, 889-902.

Brierley, I., Jenner, A. J. and Inglis, S. C. (1992). Mutational analysis of the "slippery-sequence" component of a coronavirus ribosomal frameshifting signal. *Journal of Molecular Biology* **227**, 463-479.

Buchholz, C. J., Spehner, D., Drillien, R., Neubert, W. and Homman, H. E. (1993). The conserved N-terminal region of Sendai virus nucleocapsid protein NP is required for nucleocapsid assembly. *Journal of Virology* **67**, 5803-5812.

Buckland, R. and Wild, F. (1989). Leucine zipper motif extends. *Nature* **338**, 547.

Caldwell, S. E. and Lyles, D. S. (1986). Dissociation of newly synthesized Sendai viral proteins from the cytoplasmic surface of isolated plasma membranes of infected cells. *Journal of Virology* **57**, 678-683.

Cane, P. A., Matthews, D. A. and Pringle, C. R. (1991). Identification of the variable domains of the attachment (G) protein of subgroup A respiratory syncytial viruses. *Journal of General Virology* **72**, 2091-2096.

Capone, J. P., Sedivy, J. M., Sharp, P. A. and Rajbhandary, U. L. (1986). Introduction of UAG, UAA, and UGA nonsense mutations at a specific site in the *Escherichia coli* chloramphenicol acetyltransferase gene : Use in measurement of amber, ochre, and opal suppression in mammalian cells. *Molecular and Cellular Biology* **6**, 3059-3067.

Caravokyri, C. and Pringle, C. R. (1991). Defective synthesis of envelope proteins by temperature-sensitive mutants representing complementation groups B and D of respiratory syncytial virus. *Journal of General Virology* **72**, 2501-2508.

Caravokyri, C. and Pringle, C. R. (1992). Effect of changes in the nucleotide sequence of the P gene of respiratory syncytial virus on the electrophoretic mobility of the P protein. *Virus Genes* **6**, 53-62.

Carthew, P. and Sparrow, S. (1980). Persistence of pneumonia virus of mice and Sendai virus in germ-free (nu/nu) mice. *British Journal Experimental Pathology* **61**, 172-175.

Cash, P., Wunner, W. H. and Pringle, C. R. (1977). A comparison of the polypeptides of human and bovine respiratory syncytial virus and murine pneumonia virus. *Virology* **82**, 369-379.

Cash, P., Preston, C. M. and Pringle, C. R. (1979). Characterization of murine pneumonia virus proteins. *Virology* **96**, 442-452.

Cattaneo, R., Schmid, A., Rebmann, G. and Baczko, K. (1986). Accumulated measles virus mutations in a case of subacute sclerosing panencephalitis : interrupted matrix protein reading frame and transcription alteration. *Virology* **154**, 97-107.

Cattaneo, R., Kaelin, K., Baczko, K. and Billeter, M. A. (1989). Measles virus editing provides an additional cysteine-rich protein. *Cell* **56**, 759-764.

Chakrabarti, S., Brechling, K. and Moss, B. (1985). Vaccinia virus expression vector : Coexpression of B-galactosidase provides visual screening of recombinant virus plaques. *Molecular and Cellular Biology* **5**, 3403-3409.

Chambers, P., Barr, J., Pringle, C. R. and Easton, A. J. (1990a). Molecular cloning of pneumonia virus of mice. *Journal of Virology* **64**, 1869-1872.

Chambers, P., Matthews, D. A., Pringle, C. R. and Easton, A. J. (1990b). The nucleotide sequences of intergenic regions between nine genes of pneumonia virus of mice establish the physical order of these genes in the viral genome. *Virus Research* **18**, 263-270.

Chambers, P., Pringle, C. R. and Easton, A. J. (1990c). Heptad repeat sequences are located adjacent to hydrophobic regions in several types of virus fusion glycoproteins. *Journal of General Virology* **71**, 3075-3080.

Chambers, P., Pringle, C. R. and Easton, A. J. (1991). Genes 1 and 2 of pneumonia virus of mice encode proteins which have little homology with the 1C and 1B proteins of human respiratory syncytial virus. *Journal of General Virology* **72**, 2545-2550.

Chambers, P., Pringle, C. R. and Easton, A. J. (1992). Sequence analysis of the gene encoding the fusion glycoprotein of pneumonia virus of mice suggests possible conserved secondary structure elements in paramyxovirus fusion glyoproteins. *Journal of General Virology* **73**, 1717-1724.

Chattopadhyay, D. and Banerjee, A. K. (1987). Phosphorylation within a specific domain of the phosphoprotein of Vesicular stomatitis virus regulates transcription in vitro. *Cell* **49**, 407-414.

Chen, C. and Okayama, H. (1987). High-efficiency transformation of mammalian cells by plsamid DNA. *Molecular and Cellular Biology* **7**, 2745-2752.

Chen, S. S. L. and Huang, A. S. (1986). Further characterization of the Vesicular stomatitis virus temperature-sensitive 045 mutant : intracellular conversion of the glycoprotein to a soluble form. *Journal of Virology* **59**, 210-215.

Cherrie, A. H., Anderson, K., Wertz, G. W. and Openshaw, P. J. M. (1992). Human cytotoxic T cells stimulated by antigen on dendritic cells recognize the N, SH, F, M, 22K, and 1b proteins of respiratory syncytial virus. *Journal of Virology* **66**, 2102-2110.

Chong, L. D. and Rose, J. K. (1993). Membrane association of functional vesicular stomatitis virus matrix protein in vivo. *Journal of Virology* **67**, 407-414.

Clare, J. J., Belcourt, M. and Farabaugh, P. J. (1988). Efficient translational frameshifting occurs within a conserved sequence of the overlap between the two genes of a yeast Ty1 transposon. *Proceedings of the National Academy, USA* **85**, 6816-6820.

Cohen, S. N., Chang, A. C. Y. and Hsu, L. (1972). Nonchromosomal antibiotic resistance in bacteria : Genetic trsnaformation of *Escherichia coli* by R-factor DNA. *Proceedings of the National Academy of Sciences, USA* **69**, 2110-2114.

Collier, N. C., Knox, K. and Schlesinger, M. J. (1991). Inhibition of influenza virus formation by a peptide that corresponds to sequences in the cytoplasmic domain of the hemagglutinin. *Virology* **183**, 769-772.

Collins, P. L. (1990). O glycosylation of glycoprotein G of human respiratory syncytial virus is specified within the divergent ectodomain. *Journal of Virology* **64**, 4007-4012.

Collins, P. L. (1991) The molecular biology of human respiratory syncytial virus (RSV) of the genus pneumovirus in *The Paramyxoviruses*, ed. Kingsbury, D. W. (Plenum Press, New York and London), pp. 103-162.

Collins, P. L. and Mottet, G. (1992). Oligomerization and post-translational processing of glycoprotein G of human respiratory syncytial virus: Altered O-glycosylation in the presence of brefeldin A. *Journal of General Virology* **73**, 849-863.

Collins, P. L. and Mottet, G. (1993). Membrane orientation and oligomerization of the small hydrophobic protein of human respiratory syncytial virus. *Journal of General Virology* **74**, 1445-1450.

Collins, P. L., Huang, Y. T. and Wertz, G. W. (1984). Nucleotide sequence of the gene encoding the fusion F glycoprotein of human respiratory syncytial virus. *Proceedings of the National Academy of Sciences, USA* **81**, 7683-7687.

Collins, P. L., Olmsted, R. A. and Johnson, P. R. (1990). The small hydrophobic protein of human respiratory syncytial virus: comparison between antigenic subgroups A and B. *Journal of General Virology* **71**, 1571-1576.

Collins, P. L., Mink, M. A. and Stec, D. S. (1991). Rescue of synthetic analogs of respiratory syncytial virus genomic RNA and effect of truncations and

mutations on the expression of a foreign reporter gene. *Proceedings of the National Academy of Sciences, USA* **88**, 9663-9667.

Compans, R. W., Harter, D. H. and Choppin, P. W. (1967). Studies of pneumonia virus of mice (PVM) in cell culture. II. Structure and morphogenesis of the virus particle. *Journal of Experimental Medicine* **126**, 275-277.

Connors, M., Collins, P. L., Firestone, C. Y. and Murphy, B. R. (1991). Respiratory syncytial virus (RSV) F, G, M2 (22K), and N proteins each induce resistance to RSV challenge, but resistance induced by M2 and N proteins is relatively short-lived. *Journal of Virology* **65**, 1634.

Connors, M., Kulkarni, A. B., Collins, P. L., Firestone, C. Y., Holmes, K. L., Morse, H. C., III and Murphy, B. R. (1992). Resistance to respiratory syncytial virus (RSV) challenge induced by infection with a vaccinia virus recombinant expressing the RSV M2 protein (Vac-M2) is mediated by CD8-positive T cells, while that induced by Vac-F or Vac-G recombinants is mediated by antibodies. *Journal of Virology* **66**, 1277-1281.

Coulon, P., Deutsch, V., Lafay, F., Martinet-Edelist, C., Wyers, F., Herman, R. C. and Flammand, A. (1990). Genetic evidence for multiple functions of the matrix protein of vesicular stonatitis virus. *Journal of General Virology* **71**, 991-996.

Curran, J. and Kolakofsky, D. (1988). Ribosomal initiation from an ACG codon in the Sendai virus P/C mRNA. *EMBO Journal* **7**, 245-251.

Curran, J., Boeck, R. and Kolakofsky, D. (1991). The Sendai virus P gene expresses both an essential protein and an inhibitor of RNA synthesis by shuffling modules via mRNA editing. *EMBO Journal* **10**, 3079-3085.

Curran, J., Marq, J. B. and Kolakofsky, D. (1992). The Sendai virus nonstructural C proteins specifically inhibit viral mRNA synthesis. *Virology* **189**, 647-656.

den Boon, J. A., Snijder, E. J., Chirnside, E. D., de Vries, A. A. F., Horzinek, M. C. and Spaan, W. J. M. (1991). Equine arteritis virus is not a togavirus but

belongs to the coronaviruslike superfamily. *Journal of Virology* **65**, 2910-2920.

Deshpande, K. L. and Portner, A. (1985). Monoclonal antibodies to the P protein of Sendai virus define its structure and role in transcription. *Virology* **140**, 125-134.

Dickens, L. E., Collins, P. L. and Wertz, G. W. (1984). Transcriptional mapping of human respiratory syncytial virus. *Journal of Virology* **52**, 364-369.

Eaton, M. D. and van Herick, W. (1944). Demonstration in cotton rats and rabbits of a latent virus related to pneumonia virus of mice. *Proceedings of the Society for Experimental Biology and Medicine* **57**, 89-92.

Ebata, S. N., Cote, M., Kang, Y. and Dimock, K. (1991). The fusion and hemagglutinin-neuramindase glycoproteins of human parainfluenza virus are both required for fusion. *Virology* **183**, 437-441.

Elango, N., Satake, M., Coligan, J. E., Norrby, E., Camargo, E. and Venkatesan, S. (1985). Respiratory syncytial virus fusion glycoprotein: nucleotide sequence of mRNA, identification of cleavage activation site and amino acid sequence of N-terminus of F1 subunit. *Nucleic Acids Research* **13**, 1559-1574.

Elango, N., Coligan, J. E., Jambou, R. C. and Venkatesan, S. (1986a). Human parainfluenza type 3 virus hemagglutinin-neuraminidase glycoprotein : nucleotide sequence of mRNA and limited amino acid sequence of the purified protein. *Journal of Virology* **57**, 481-489.

Elango, N., Prince, G. A., Murphy, B. R., Venkatesan, S., Chanock, R. M. and Moss, B. (1986b). Resistance to human respiratory syncytial virus (RSV) infection induced by immunization of cotton rats with a recombinant vaccinia virus expressing the RSV G glycoprotein. *Proceedings of the National Academy of Sciences, USA* **83**, 1906-1910.

Elango, N., Kovamees, J., Varsanyi, T. M. and Norrby, E. (1989). mRNA sequence and deduced amino acid sequence of the mumps virus small hydrophobic protein gene. *Journal of Virology* **64**, 4137-4145.

Faaberg, K. S. and Peeples, M. E. (1988). Association of soluble matrix protein of Newcastle disease virus with liposomes is independent of ionic conditions. *Virology* **166**, 123-132.

Felsenstein, K. M. and Goff, S. P. (1988). Expression of the gag-pol fusion protein of Moloney murine leukemia virus without gag protein does not induce virion formation or proteolytic processing. *Journal of Virology* **62**, 2179-2182.

Feng, Y. X., Levin, J. G., Hatfield, D. L., Schaefer, T. S., Gorelick, R. J. and Rein, A. (1989). Suppression of UAA and UGA termination codons in mutant murine leukemia viruses. *Journal of Virology* **63**, 2870-2873.

Fernie, B. F., Dapolito, G., Cote, P. J. and Gerin, J. L. (1985). Kinetics of synthesis of RS virus glycoproteins. *Journal of General Virology* **66**, 1983-1990.

Flower, A. M. and McHenry, C. S. (1990). The gamma subunit of DNA polymerase III holoenzyme of *Escherichia coli* is produced by ribosomal frameshifting. *Proceedings of the National Academy of Sciences, USA* **87**, 3713-3717.

Fooks, A., Stephenson, J. R., Warnes, A., Dowsett, A. B., Rima, B. K. and Wilkinson, G. W. G. (1993). Measles virus nucleocapsid protein expressed in insect cells assembles into nucleocapsid-like structures. *Journal of General Virology* **74**, 1439-1444.

Francki, R. I. B., Fauquet, C. M., Knudson, D. L. and Brown, F. (1991) Classification and nomenclature of viruses. Fifth report of the international committee on taxonomy of viruses. Archives of Virology. Springer-Verlag, Wein, New York.

Fuerst, T. R., Niles, E. G., Studier, W. and Moss, B. (1986). Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes

bacteriophage T7 RNA polymerase. *Proceedings of the National Academy of Sciences, USA* **83**, 8122-8126.

Galinski, M. S., Troy, R. M. and Banerjee, A. K. (1992). RNA editing in the phosphoprotein gene of the human parainfluenza virus type 3. *Virology* **186**, 543-550.

Gannon, J. and Carthew, P. (1980). Prevalence of indigenous viruses in laboratory animal colonies in the United Kingdom 1978-1979. *Laboratory Animals* **14**, 309-311.

Garcia, J., Garcia-Barreno, B., Vivo, A. and Melero, J. A. (1993). Cytoplasmic inclusions of respiratory syncytial virus-infected cells: Formation of inclusion bodies in transfected cells that coexpress the nucleoprotein, the phosphoprotein, and the 22K protein. *Virology* **195**, 243-247.

Garcia-Barreno, B., Palomo, C., Penas, C., Delgado, C., Perez-Brena, P. and Melero, J. (1989). Marked differences in the antigenic structure of human respiratory syncytial virus F and G glycoproteins. *Journal of Virology* **63**, 925-932.

Garcia-Barreno, B., Portela, A., Delgado, T. and Lopez, J. (1990). Frame shift mutations as a novel mechanism for the generation of neutralization resistant mutants of human respiratory syncytial virus. *EMBO Journal* **12**, 4181-4187.

Garcia-Barreno, B., Delgado, T., Akerlind-Stopner, B., Norrby, E. and Melero, J. A. (1992). Location of the epitope recognized by monoclonal antibody 63G on the primary structure of human respiratory syncytial virus G glycoprotein and the ability of synthetic peptides containing this epitope to induce neutralizing antibodies. *Journal of General Virology* **73**, 2625-2630.

Gill, D. S., Chattopadhyay, D. and Banerjee, A. K. (1986). Identification of a domain within the phosphoprotein of vesicular stomatitis virus that is essential for transcription *in vitro*. *Proceedings of the National Academy of Sciences, USA* **83**, 8873-8877.

Gimenez, H. B., Cash, P. and Melvin, W. T. (1984). Monoclonal antibodies to human respiratory syncytial virus and their use in comparison of different isolates. *Journal of General Virology* **65**, 963-971.

Glickman, R. L., Syddall, R. J., Iorio, R. M., Sheehan, J. P. and Bratt, M. A. (1988). Quantitative basic residue requirements in the cleavage-activation site of the fusion glycoprotein as a determinant of virulence for Newcastle disease virus. *Journal of Virology* **62**, 354-356.

Gombart, A. F., Hirano, A. and Wong, T. C. (1993). Conformational maturation of measles virus nucleocapsid protein. *Journal of Virology* **67**, 4133-4141.

Gupta, K. C. and Patwardhan, S. (1988). ACG, the initiator codon for a Sendai virus protein. *Journal of Biological Chemistry* **263**, 8553-8556.

Hamaguchi, M., Yoshida, T., Nishikawa, K., Naruse, H. and Nagai, Y. (1983). Transcriptive complex of Newcastle disease virus. I. Both L and P proteins are required to constitute an active complex. *Virology* **128**, 105-117.

Harter, D. H. and Choppin, P. W. (1967). Studies on pneumonia virus of mice (PVM) in cell culture. I. Replication in baby hamster kidney cells and properties of the virus. *Journal of Experimental Medicine* **126**, 251-266.

Hendricks, D. A., Baradaran, K., McIntosh, K. and Patterson, J. L. (1987). Appearance of a soluble form of the G protein of respiratory syncytial virus in fluids of infected cells. *Journal of General Virology* **68**, 1705-1714.

Hendricks, D. A., McIntosh, K. and Patterson, J. L. (1988). Further characterization of the soluble form of the G glycoprotein of respiratory syncytial virus. *Journal of Virology* **62**, 2228-2233.

Hiebert, S. W., Paterson, R. G. and Lamb, R. A. (1985). Hemagglutinin-neuraminidase protein of the paramyxovirus simian virus 5 : nucleotide sequence of the mRNA predicts an N-terminal membrane anchor. *Journal of Virology* **54**, 1-6.

Hiebert, S. W., Richardson, C. D. and Lamb, R. A. (1988). Cell surface expression and orientation in membranes of the 44-amino-acid SH protein in simian virus 5. *Journal of Virology* **62**, 2347-2357.

Hirano, A., Wang, A. H., Gombart, A. F. and Wong, T. C. (1992). The matrix proteins of neurovirulent subacute sclerosing panencephalitis virus and its acute measles virus progenitor are functionally different. *Proceedings of the National Academy of Sciences, USA* **89**, 8745-8749.

Homann, H. E., Willenbrink, W., Buchholz, C. J. and Neubert, W. J. (1991). Sendai virus protein-protein interactions studied by a protein-blotting protein-overlay technique: Mapping of domains on NP protein required for binding to P protein. *Journal of Virology* **65**, 1304-1309.

Hopp, T. J. and Woods, K. L. (1981). Prediction of protein antigenic determinants from amino acid sequences. *Proceedings of the National Academy of Sciences, USA* **78**, 3824-3828.

Horikami, S. M., Curran, J., Kolakofsky, D. and Moyer, S. A. (1992). Complexes of Sendai virus NP-P and P-L proteins are required for defective interfering particle genome replication in vitro. *Journal of Virology* **66**, 4901-4908.

Horsfall, F. L. and Curnen, E. C. (1946). Studies on pneumonia virus of mice (PVM). II. Immunological evidence of latent infection with the virus in numerous mammalian species. *Journal of Experimental Medicine* **83**, 43-64.

Horsfall, F. L. and Ginsberg, H. S. (1951). The dependence of the pathological lesion upon the multiplication of pneumonia virus of mice (PVM). Kinetic relation between the degree of viral multiplication and the extent of pneumonia. *Journal of Experimental Medicine* **93**, 151-160.

Horsfall, F. L. and Hahn, R. G. (1939). A pneumonia virus of Swiss mice. *Proceedings of the Society for Experimental Biology and Medicine* **40**, 684-686.

Horsfall, F. L. and Hahn, R. G. (1940). A latent virus in normal mice capable of producing pneumonia in its natural host. *Journal of Experimental Medicine* **71**, 391-408.

Horsfall, F. L., Curnen, E. C., Mirick, G. S., Thomas, L. and Ziegler Jr., J. E. (1943). A virus recovered from patients with primary atypical pneumonia. *Science* **97**, 289-291.

Horvath, C. M. and Lamb, R. A. (1992). Studies on the fusion peptide of a paramyxovirus fusion glycoprotein: Roles of conserved residues in cell fusion. *Journal of Virology* **66**, 2443-2455.

Howard, M. and Wertz, G. W. (1989). Vesicular stomatitis virus RNA replication : A role for the NS protein. *Journal of General Virology* **70**, 2683-2694.

Hsu, M.-C., Scheid, A. and Choppin, P. W. (1987). Protease activation mutants of Sendai virus : sequence analysis of the mRNA of the fusion protein (F) gene and direct identification of the cleavage-activation site. *Virology* **156**, 84-90.

Hu, X., Ray, R. and Compans, R. W. (1992). Functional interactions between the fusion protein and hemagglutinin-neuraminidase of human parainfluenza viruses. *Journal of Virology* **66**, 1528-1534.

Huang, Y. T. and Wertz, G. W. (1982). The genome of respiratory syncytial virus is a negative-stranded RNA that codes for at least seven mRNA species. *Journal of Virology* **43**, 150-157.

Huang, Y. T., Collins, P. L. and Wertz, G. (1984) Identification of a new envelope-associated protein of human respiratory syncytial virus in *Nonsegmented negative strand viruses, Paramyxoviruses and Rhabdoviruses*, eds. Bishop, D. H. L. and Compans, R. W. (Acedemic Press, New York), pp. 365-368.

Huang, Y. T., Collins, P. L. and Wertz, G. W. (1985). Characterization of the 10 proteins of human respiratory syncytial virus: Identification of a fourth envelope-associated protein. *Virus Research* **2**, 157-173.

Huber, M., Cattaneo, R., Spielhofer, P., Orvell, C., Norrby, E., Messerli, M., Perriard, J. C. and Billeter, M. A. (1991). Measles virus phosphoprotein retains the nucleocapsid protein in the cytoplasm. *Virology* **185**, 299-308.

Ishida, N. and Homma, N. (1978). Sendai virus. *Advances in Virus Research* **23**, 349-383.

Itoh, M. and Homma, H. (1988). Single amino acid change at the cleavage site of the fusion protein is responsible for both enhanced chymotrypsin sensitivity and trypsin resistance of a Sendai virus mutant TR-5. *Journal of General Virology* **69**, 2907-2911.

Itoh, M., Shibuta, H. and Homma, M. (1987). Single amino acid substitution of Sendai virus at the cleavage site of the fusion protein confers trypsin resistance. *Journal of General Virology* **68**, 2939-2944.

Jacks, T. and Varmus, H. E. (1985). Expression of the Rous sarcoma virus *pol* gene by ribosomal frameshifting. *Science* **230**, 1237-1242.

Jacks, T., Townsley, K., Varmus, H. E. and Majors, J. (1987). Two efficient ribosomal frameshifting events are required for synthesis of mouse mammary tumor virus *gag*-related polyproteins. *Proceedings of the National Academy of Sciences, USA* **84**, 4298-4302.

Jacks, T., Madhani, H. D., Masiarz, F. R. and Varmus, H. E. (1988). Signals for ribosomal frameshifting in the Rous sarcoma virus gag-pol region. *Cell* **55**, 447-458.

Johnson, P. R. and Collins, P. L. (1988). The fusion glycoproteins of human respiratory syncytial virus of subgroups A and B : sequence conservation provides a structural basis for antigenic relatedness. *Journal of General Virology* **69**, 2623-2628.

Johnson, P. R., Olmsted, R. A., Prince, G. A., Murphy, B. R., Alling, D. W., Walsh, E. E. and Collins, P. L. (1987a). Antigenic relatedness between glycoproteins of

human respiratory suncytial virus subgroups A and B: Evaluation of the contributions of F and G glycoproteins to immunity. *Journal of Virology* **61**, 3163-3166.

Johnson, P. R., Spriggs, M. K., Olmsted, R. A. and Collins, P. L. (1987b). The G glycoprotein of human respiratory syncytial viruses of subgroups A and B : Extensive sequence divergence between antigenically related proteins. *Proceedings of the National Academy of Sciences, USA* **84**, 5625-5629.

Jorgensen, E. D., Collins, P. L. and Lomedico, P. T. (1987). Cloning and nucleotide sequence of Newcastle disease virus hemagglutinin-neuraminidase mRNA : Identification of a putative sialic acid binding site. *Virology* **156**, 12-24.

Kim, H. W., Canchola, J. G., Brandt, C. D., Pyles, G., Chanock, R. M., Jensen, K. and Parrott, R. H. (1969). Respiratory syncytial virus disease in infants despite prior administration of antigenic inactivated vaccine. *American Journal of Epidemiology* **89**, 422-434.

Kingsbury, D. W., Bratt, M. A., Choppin, P. W., Hanson, R. P., Hosaka, Y., Meulen, T., Norrby, E., Plowright, W., Rott, R. and Wunner, W. H. (1978). Paramyxoviridae. *Intervirology* **10**, 137-152.

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283-292.

Kozak, M. (1989). The scanning model for translation : an update. *Journal of Cell Biology* **108**, 229-241.

Kumar, A. and Lindberg, U. (1972). Characterisation of messenger ribonucleoprotein and messenger RNA from KB cells. *Proceedings of the National Academy of Sciences, USA* **69**, 681-685.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Lambert, D. M. (1988). Role of oligosaccharides in the structure and function of respiratory syncytial virus glycoproteins. *Virology* **164**, 458-466.

Lambert, D. M. and Pons, M. W. (1983). Respiratory syncytial virus glycoproteins. *Virology* **130**, 204-214.

Levine, S., Klaiber-Franco, R. and Paradiso, P. R. (1987). Demonstration that glycoprotein G is the attachment protein of respiratory syncytial virus. *Journal of General Virology* **68**, 2521-2524.

Li, Y., Schubert, M., Wagner, R. R. and Kang, C. Y. (1993). Viral liposomes released from insect cells infected with recombinant baculovirus expressing the matrix protein of vesicular stomatitis virus. *Journal of Virology* **67**, 4415-4420.

Lichtenstein, D. L., Wertz, G. W. and Ball, L. A. (1991) Mutational analysis of the signal/anchor domain of the glycoprotein G of human respiratory syncytial virus.

Ling, R. and Pringle, C. R. (1989a). Polypeptides of pneumonia virus of mice. I: Immunological cross-reactions and post-translational modifications. *Journal of General Virology* **70**, 1427-1440.

Ling, R. and Pringle, C. R. (1989b). Polypeptides of pneumonia virus of mice. II: Characterization of the glycoproteins. *Journal of General Virology* **70**, 1441-1452.

Ling, R., Easton, A. J. and Pringle, C. R. (1992). Sequence analysis of the 22K, SH and G genes of turkey rhinotracheitis virus and their intergenic regions reveals a gene order different from that of other pneumoviruses. *Journal of General Virology* **73**, 1709-1715.

Lopez, J. A., Penas, C., Garcia-Barreno, B., Melero, J. A. and Portela, A. (1990). Location of a highly conserved neutralizing epitope in the F glycoprotein of human respiratory syncytial virus. *Virology* **64**, 927-930.

Luytjes, W., Krystal, M., Enami, M., Parvin, J. D. and Palese, P. (1989). Amplification, expression, and packaging of a foreign gene by influenza virus. *Cell* **59**, 1107-1113.

Lyles, D. S., Puddington, L. and McCreedy, B. J. (1988). Vesicular stomatitis virus M protein in the nuclei of infected cells. *Journal of Virology* **62**, 4387-4392.

Martin-Gallardo, A., Fien, K. A., Hu, B. T., Farley, J. F., Seid, R., Collins, P. L., Hildreth, S. W. and Paradiso, P. R. (1991). Expression of the F glycoprotein gene from human respiratory syncytial virus in *Escherichia coli*: mapping of a fusion inhibiting epitope. *Virology* **184**, 428-432.

Martin-Gallardo, A., Fleischer, E., Doyle, S. A., Arumugham, R., Collins, P. L., Hildreth, S. W. and Paradiso, P. R. (1993). Expression of the G glycoprotein of human respiratory syncytial virus in Salmonella typhimurium. *Journal of General Virology* **74**, 453-458.

Masters, P. S. and Banerjee, A. K. (1988a). Resolution of multiple complexes of phosphoprotein NS with nucleocapsid protein N of vesicular stomatitis virus. *Journal of Virology* **62**, 2651-2657.

Masters, P. S. and Banerjee, A. K. (1988b). Complex formation with vesicular stomatitis virus phosphoprotein NS prevents binding of nucleocapsid protein N to non-specific RNA. *Journal of Virology* **62**, 2658-2664.

Matsuoka, Y., Curran, J., Pelet, T., Kolakofsky, D., Ray, R. and Compans, R. W. (1991). The P gene of human parainfluenza virus type 1 encodes P and C proteins but not a cysteine-rich V protein. *Journal of Virology* **65**, 3406-3410.

McGinnes, L. W., McQuain, C. and Morrison, T. (1988). The P protein and the non-structural 38K and 29K proteins of Newcastle disease virus are derived from the same open reading frame. *Virology* **164**, 256-264.

Mellor, J., Fulton, S. M., Dobson, M. J., Wilson, W., Kingsman, S. M. and Kingsman, A. J. (1985). A retrovirus-like strategy for expression of a fusion protein by yeast transposon Tyl. *Nature* **313**, 243-246.

Metsikko, K. and Simons, K. (1986). The budding mechanism of spikeless vesicular stomatitis virus particles. *EMBO Journal* **5**, 1913-1920.

Millar, N. S., Chambers, P. and Emmerson, P. T. (1986). Nucleotide sequence analysis of the haemaglutinin-neuraminidase gene of Newcastle disease virus. *Journal of General Virology* **67**, 1917-1927.

Mills, K. C. and Dochez, A. R. (1944). Specific agglutination of murine erythrocytes by a pneumonitis virus in mice. *Proceedings of the Society for Experimental Biology and Medicine* **57**, 140.

Mink, M. A., Stec, D. S. and Collins, P. L. (1991). Nucleotide sequences of the 3' leader and 5' trailer regions of human respiratory syncytial virus genomic RNA. *Virology* **185**, 615-624.

Moore, R., Dixon, M., Smith, R., Peters, G. and Dickson, C. (1987). Complete nucletide sequence of a milk-transmitted mouse mammary tumor virus: two frameshifting suppression events required for translation of *gag* and *pol*. *Journal of Virology* **61**, 480-490.

Morikawa, S. and Bishop, D. H. L. (1992). Identification and analysis of the gag-pol ribosomal frameshift site of feline immunodeficiency virus. *Virology* **186**, 389-397.

Morikawa, S., Booth, T. F. and Bishop, D. H. L. (1991). Analyses of the requirements for the synthesis of virus-like particles by feline immunodeficiency virus gag using baculovirus vectors. *Virology* **183**, 288-297.

Morimoto, K., Iwatani, Y. and Kawai, A. (1993). Shedding of Gs protein (a soluble from of the viral glycoprotein) by the rabies virus-infected BHK-21 cells. *Virology* **195**, 541-549.

Morrison, T. G. (1988). Structure, function, and intracellular processing of paramyxovirus membrane proteins. *Virus Research* **10**, 113-136.

Morrison, T. and Portner, A. (1991) Structure, function, and intracellular processing of the glycoproteins of *Paramyxoviridae* in *The Paramyxoviruses*, ed. Kingsbury, D. W. (Plenum Press, New York and London), pp. 347-382.

Morrison, T., McQuain, C. and McGinnes, L. (1991). Complementation between avirulent Newcastle disease virus and a fusion protein gene expressed from a retrovirus vevtor: Requirements for membrane fusion. *Journal of Virology* **65**, 813-822.

Morrison, T., McQuain, C., Sergel, T., McGinnes, L. and Reitter, J. (1993). The role of the amino terminusof F1 of the Newcastle disease virus fusion protein in cleavage and fusion. *Virology* **193**, 997-1000.

Mufson, M. A., Orvell, C., Rafnar, B. and Norrby, E. (1985). Two distinct subtypes of respiratory syncytial virus. *Journal of General Virology* **66**, 2111-2124.

Muramatsu, M. and Homma, M. (1980). Trypsin action on the growth of Sendai virus in tissue culture cells. V. An activating enzyme for Sendai virus in the chorioallantoic fluid of the embryonated chicken ehh. *Microbiology and Immunology* **24**, 113-122.

Nagai, Y. and Klenk, H.-D. (1977). Activation of precursors to both precursors of Newcastle disease virus by proteolytic cleavage. *Virology* **77**, 125-134.

Nagai, Y., Klenk, H.-D. and Rott, R. (1976). Proteolytic cleavage of the viral glycoproteins and its significance for the virulence of Newcastle disease virus. *Virology* **72**, 494-508.

Nagy, E., Huber, P., Krell, P. J. and Derbyshire, J. B. (1991). Synthesis of Newcastle Disease Virus (NDV)-like envelopes in insect cells infected with a recombinant baculovirus expressing the haemagglutinin-neuraminidase of NDV. *Journal of General Virology* **72**, 753-756.

Nam, S. H., Copeland, T. D., Hatanaka, M. and Oroszlan, S. (1993). Characterization of ribosomal frameshifting for expression of pol gene products of human T-cell leukemia virus type I. *Journal of Virology* **67**, 196-203.

Nicholas, J. A., Rubino, K. L., Levely, M. E., Meyer, A. L. and Collins, P. L. (1991). Cytotoxic T cell activity against the 22-kDa protein of human respiratory syncytial virus (RSV) is associated with a significant reduction in pulmonary RSV replication. *Virology* **182**, 664-672.

Ogasawara, T., Gotoh, B., Suzuki, H., Asaka, J. I., Shimokata, K., Rott, R. and Nagai, Y. (1992). Expression of factor X and its significance for the determination of paramyxovirus tropism in the chick embryo. *EMBO Journal* **11**, 467-472.

Ogden, J. R., Pal, R. and Wagner, R. R. (1986). Mapping regions of the matrix protein of vesicular stomatitis virus which bind to ribonucleocapsids, liposomes and monoclonal antibodies. *Journal of Virology* **58**, 860-868.

Ohgimoto, S., Bando, H., Kawano, M., Okamoto, K., Kondo, K., Tsurudome, M., Nishio, m and Ito, Y. (1990). Sequence analysis of P gene of human parainfluenza type 2 virus: P and cysteine-rich proteins are translated by two mRNAs that differ by two nontemplated G residues. *Virology* **177**, 116-123.

Olmsted, R. A. and Collins, P. L. (1989). The 1A protein of respiratory syncytial virus is an integral membrane protein present as multiple, structurally distinct species. *Journal of Virology* **63**, 2019-2029.

Olmsted, R. A., Elango, N., Prince, G. A., Murphy, B. R., Johnson, P. R., Moss, B., Chanock, R. M. and Collins, P. L. (1986). Expression of the F glycoprotein of respiratory syncytial virus by a recombinant vaccinia virus: Comparison of the individual contributions of the F and G glycoproteins to host immunity. *Proceedings of the National Academy of Sciences, USA* **83**, 7462-7466.

Olmsted, R. A., Murphy, B. R., Lawrence, L. A., Elango, N., Moss, B. and Collins, P. L. (1989). Processing, surface expression, and immunogenicity of carboxy-terminally truncated mutants of G protein of human respiratory syncytial virus. *Journal of Virology* **63**, 411-420.

Owens, R. J. and Rose, J. K. (1993). Cytoplasmic domain requirement for incorporation of a foreign envelope protein into vesicular stomatitis virus. *Journal of Virology* **67**, 360-365.

Palomo, C., Garcia-Barreno, B., Penas, P. and Melero, J. A. (1991). The G protein of hyman respiratory syncytial virus: significance of carbohydrate side chains and the C-terminal end to its antigenicity. *Journal of General Virology* **72**, 669-675.

Park, J. and Morrow, C. D. (1991). Overexpression of the gag-pol precursor from human immunodeficiency virus type 1 proviral genomes results in efficient proteolytic processing in the absence of virion production. *Journal of Virology* **65**, 5111-5117.

Park, K. H., Huang, T. H., Correia, F. F. and Krystal, M. (1991). Rescue of a foreign gene by Sendai virus. *Proceedings of the National Academy of Sciences, USA* **88**, 5537-5541.

Parkin, N. T., Chamorro, M. and Varmus, H. E. (1992). Human immunodeficiency virus type 1 gag-pol frameshifting is dependent on downstream mRNA secondary structure: Demonstration by expression in vivo. *Journal of Virology* **66**, 5147-5151.

Paterson, R. G. and Lamb, R. A. (1990). RNA editing by nucleotide insertion in mumps virus P-gene mRNA transcripts. *Journal of Virology* **64**, 4137-4145.

Pattnaik, A. and Wertz, G. W. (1990). Replication and amplification of defective interfering particle RNAs of vesicular stomatitis virus in cells expressing viral

proteins from vectors containing cloned cDNAs. *Journal of Virology* **64**, 2948-2957.

Pattnaik, A. and Wertz, G. W. (1991). Cells that express all five proteins of vesicular stomatitis virus from cloned cDNAs support replication, assembly, and budding of defective interfering particles. *Proceedings of the National Academy of Sciences, USA* **88**, 1379-1383.

Peeples, M. E. (1991) Paramyxovirus M proteins: Pulling it all together and taking it on the road in *The Paramyxoviruses*, ed. Kingsbury, D. W. (Plenum Press, New York and London), pp. 427-456.

Peeples, M. E. and Bratt, M. A. (1984). Mutation in the matrix protein of Newcastle disease virus can result in decreased fusion glycoprotein incorporation into particles and decreased infectivity. *Journal of Virology* **51**, 81-90.

Peeples, M. E., Wang, C., Gupta, K. C. and Coleman, N. (1992). Nuclear entry and nuclear localization of the Newcastle disease virus (NDV) matrix protein occur early in infection and do not require other NDV proteins. *Journal of Virology* **66**, 3263-3269.

Power, U. F., Ryan, K. W. and Portner, A. (1992). The P genes of human parainfluenza virus type 1 clinical isolates are polycistronic and microheterogeneous. *Virology* **189**, 340-343.

Pringle, C. R. (1987) Paramyxoviruses and disease in *Molecular basis of virus disease, Society for general microbiology sysposium*, eds. Russell, W. C. and Almond, J. W. (Cambridge University Press, Cambridge), Vol. 40, pp. 91-138.

Pringle, C. R. and Eglin, R. P. (1986). Murine pneumonia virus: Seroepidemiological evidence of widespread human infection. *Journal of General Virology* **67**, 975-982.

Prufer, D., Tacke, E., Schmitz, J., Kull, B., Kaufmann, A. and Rohde, W. (1992). Ribosomal frameshifting in plants: A novel signal directs the -1 frameshift in

the synthesis of the putative viral replicase of potato leafroll luteovirus. *EMBO Journal* **11**, 1111-1117.

Ray, R., Roux, L. and Compans, R. W. (1991) Intracellular targeting and assembly of paramyxovirus proteins in *The Paramyxoviruses*, ed. Kingsbury, D. W. (Plenum Press, New York and London), pp. 457-479.

Reil, H., Kollmus, H., Weidle, U. H. and Hauser, H. (1993). A heptanucleotide sequence mediates ribosomal frameshifting in mammalian cells. *Journal of Virology* **67**, 5579-5584.

Richardson, C. D. and Choppin, P. W. (1983). Oligopeptides that specifically inhibit membrane fusion by paramyxoviruses : studies on the site of action. *Virology* **131**, 518-532.

Richardson, C., Hull, D., Greer, P., Hasel, K., Berkovitch, A., Englund, G., Bellini, W., Rima, B. and Lazzarini, R. (1986). The nucleotide sequence of the mRNA encoding the fusion protein of measles virus (Edmonston strain): a comparison of fusion proteins from several different paramyxoviruses. *Virology* **155**, 508-523.

Roberts, S. R., Lichtenstein, D., Ball, L. A. and Wertz, G. W. (1993) Secreted and anchored froms of the attachment protein of respiratory syncytial virus are synthesized from different translational initiation sites (on the same mRNA). Presented at the IXth International Congress of Virology, Glasgow, Scotland, U.K.

Routledge, E. G., Willcocks, M. M., Morgan, L., Samson, A. C. R., Scott, R. and Toms, G. L. (1987). Expression of the respiratory syncytial virus 22K protein on the surface of infected HeLa cells. *Journal of General Virology* **68**, 1217-1222.

Routledge, E. G., Willcocks, M. M., Samson, A. C. R., Morgan, L., Scott, R., Anderson, J. and Toms, G. L. (1988). The purification of four RS virus

proteins and their evaluation as protective agents against experimental infection in BALB/C mice. *Journal of General Virology* **69**, 293-303.

Rueda, P., Delgardo, T., Portella, A., Melero, J. A. and Garcia-Barreno, B. (1991). Premature stop codons in the G glycoprotein of human respiratory syncytial viruses resistant to neutralization by monoclonal antibodies. *Journal of Virology* **65**, 3374-3378.

Ryan, K. W. and Kingsbury, D. W. (1988). Carboxyl-terminal region of Sendai virus P protein is required for binding to viral nucleocapsids. *Virology* **167**, 106-112.

Samal, S. K., Pastey, M. K., McPhilips, T. H. and Mohanty, S. B. (1993). Bovine respiratory syncytial virus nucleocapsid protein expressed in insect cells interacts with the phosphoprotein and the M2 protein. *Virology* **193**, 470-473.

Sanger, F., Coulsen, A. R., Barrell, B. G., Smith, A. J. H. and Roe, B. A. (1980). Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *Journal of Molecular Biology* **143**, 161-178.

Satake, M. and Venkatesan, S. (1984). Nucleotide sequence of the gene encoding the respiratory syncytial virus matrix protein. *Journal of Virology* **50**, 92-99.

Satake, M., Coligan, J. E., Elango, N., Norrby, E. and Venkatesan, S. (1985). Respiratory syncytial virus envelope glycoprotein (G) has a novel structure. *Nucleic Acids Research* **13**, 7795-7812.

Schnitzer, T. J., Dickson, C. and Weiss, R. A. (1979). Morphological and biochemical characterization of viral particles produced by the *ts*045 mutant of vesicular stomatitis virus at restrictive temperature. *Journal of Virology* **162**, 417-426.

Scopes, G. E., Watt, P. J. and Lambden, P. R. (1990). Identification of a linear epitope on the fusion glycoprotein of respiratory syncytial virus. *Journal of General Virology* **71**, 53-59.

Sergel, T., McGinnes, L. W., Peeples, M. E. and Morrison, T. G. (1993). The attachment function of the Newcastle disease virus hemagglutinin-neuraminidase protein can be separated from fusion promotion by mutataion. *Virology* **193**, 717-726.

Serghini, M. A., Ritzenthaler, C. and Pinck, L. (1989). A rapid efficient 'mini-prep' for isolation of plasmid DNA. *Nucleic Acids Research* **17**, 3604.

Smith, A. L., Carrano, V. A. and Brownstein, D. G. (1984). Response of weanling random-bred mice to infection with pneumonia virus of mice (PVM). *Laboratory Animal Science* **34**, 35-37.

Spehner, D., Kirn, A. and Drillien, R. (1991). Assembly of nucleocapsid-like structures in animal cells infected with a vaccinia virus recombinant encoding the measles virus nucleoprotein. *Journal of Virology* **65**, 6296-6300.

Stec, D. S., Hill III, M. G. and Collins, P. L. (1991). Sequence analysis of the polymerase L gene of human respiratory syncytial virus and predicted phylogeny of nonsegmented negative-strand viruses. *Virology* **183**, 273-287.

Stott, E. J., Ball, L. A., Young, K. K., Furze, J. and Wertz, G. W. (1986). Human respiratory syncytial virus glycoprotein G expressed from a recombinant vaccinia virus vector protects mice against live-virus challenge. *Journal of Virology* **60**, 607-613.

Stott, E. J., Ball, L. A., Anderson, K., Young, K. K., King, A. M. Q. and Wertz, G. W. (1987). Immune and histopathological responses in animals vaccinated with recombinant vaccinia viruses that express individual genes of human respiratory syncytial virus. *Journal of Virology* **61**, 3855-3861.

Stricker, R. and Roux, L. (1991). The major glycoprotein of Sendai virus is dispensable for efficient virus particle budding. *Journal of General Virology* **72**, 1703-1707.

Sullender, W. M. and Wertz, G. W. (1991) The unusual attachment glycoprotein of the respiratory syncytial viruses: Structure, maturation, and role in immunity in *The Paramyxoviruses*, ed. Kingsbury, D. W. (Plenum Press, New York and London), pp. 383-406.

Sullender, W. M., Anderson, K. and Wertz, G. W. (1990). The respiratory syncytial virus subgroup B attachment glycoprotein: Analysis of sequence, expression from a recombinant vector, and evaluation as an immunogen against homologous and heterologous subgroup virus challenge. *Virology* **178**, 195-203.

Sullender, W. M., Mufson, M. A., Anderson, L. J. and Wertz, G. W. (1991). Genetic diversity of the attachment protein of subgroup B respiratory syncytial viruses. *Journal of Virology* **65**, 5425-5434.

Takeuchi, K., Tanabayashi, K., Hishiyama, M., Yamada, A. and Sugiura, A. (1991). Variations of nucleotide sequences and transcription of the SH gene among mumps virus strains. *Virology* **181**, 364-366.

Tanabayashi, K., Takeuchi, K., Okazaki, K., Hishiyama, M. and Yamada, A. (1992). Expression of mumps virus glycoproteins in mammalian cells from cloned cDNAs: Both F and HN proteins are required for cell fusion. *Virology* **187**, 801-804.

Tashiro, M. and Homma, M. (1983). Evidence of proteolytic activation of Sendai virus in mouse lung. *Archives of Virology* **77**, 127-137.

Taylor, G., Stott, E. J., Bew, M., Fernie, B. F., Cote, P. J., Collins, P. L., Hughes, M. and Jebbett, J. (1984). Monoclonal antibodies protect against respiratory syncytial virus infection in mice. *Immunology* **52**, 137-142.

Thomas, S. M., Lamb, R. A. and Paterson, R. G. (1988). Two mRNAs that differ by two non-templated nucleotides encode the amino coterminal proteins P and V of the paramyxovirus SV5. *Cell* **54**, 891-902.

Toyoda, T., Sakaguchi, T., Imai, K., Inocencio, N. M., Gotoh, B., Hamaguchi, M. and Nagai, Y. (1987). Structural comparison of the cleavage-activataion site of the fusion glycoprotein between virulent and avirulent strains of Newcastle disease virus. *Virology* **158**, 242-247.

Trudel, M., Nadon, F., Seguin, C., Payment, P. and Talbot, P. (1987). Respiratory syncytial virus fusion glycoprotein: Further characterization of a major epitope involved in virus neutralization. *Canadian Journal of Microbiology* **33**, 933-938.

Trudel, M., Stott, E. J., Taylor, G., Oth, D., Mercier, G., Nadon, F., Seguin, C., Simard, C. and Lacroix, M. (1991). Synthetic peptides corresponding to the F protein of RSV stimulate murine B and T cells but fail to confer protection. *Archives of Virology* **117**, 59-71.

Tsuchihashi, Z. (1991). Translational frameshifting in the *Escherichia coli* dnaX gene in vitro. *Nucleic Acids Research* **19**, 2457-2462.

Tsuchihashi, Z. and Kornberg, A. (1990). Translational frameshifting generates the gamma subunit of DNA polymerase III holoenzyme. *Proceedings of the National Academy of Sciences, USA* **87**, 2516-2520.

Tu, C., Tzeng, T. H. and Bruenn, J. A. (1992). Ribosomal movement impeded at a pseudoknot required for frameshifting. *Proceedings of the National Academy of Sciences, USA* **89**, 8636-8640.

Tzeng, T. H., Tu, C. L. and Bruenn, J. A. (1992). Ribosomal frameshifting requires a pseudoknot in the Saccharomyces cerevisiae double-stranded RNA virus. *Journal of Virology* **66**, 999-1006.

Udem, S. A. and Cook, K. A. (1984). Isolation and characterization of measles virus intracellular nucleocapsid RNA. *Journal of Virology* **49**, 57-65.

Veit, M., Herrler, G., Schmidt, M. F. G., Rott, R. and Klenk, H. (1990). The hemagglutinating glycoproteins of influenza B and C viruses are acylated with different fatty acids. *Virology* **177**, 807-811.

Veit, M., Klenk, H.-D., Kendal, A. and Rott, R. (1991). The M2 protein of influenza A virus is acylated. *Journal of General Virology* **72**, 1461-1465.

Vidal, S., Curran, J. and Kolakofsky, D. (1990). Editing of the Sendai virus P/C mRNA by G insertion occurs during mRNA synthesis via a virus-encoded activity. *Journal of Virology* **64**, 239-246.

Vijaya, S., Elango, N., Zavala, F. and Moss, B. (1988). Transport to the cell surface of a peptide sequence attached to the truncated C-terminus of an N-terminally anchored integral membrane protein. *Molecular and Cellular Biology* **8**, 1709-1714.

Walsh, E. E., Brandriss, M. W. and Schlesinger, J. J. (1987). Immunological differences between the envelope glycoproteins of two strains of human respiratory syncytial virus. *Journal of General Virology* **68**, 2169-2176.

Wang, C., Raghu, G., Morrison, T. and Peeples, M. E. (1992). Intracellular processing of the paramyxovirus F protein: Critical role of the predicted amphipathic alpha helix adjacent to the fusion domain. *Journal of Virology* **66**, 4161-4169.

Wathen, M. W., Brideau, R. J. and Thomsen, D. R. (1989a). Immunization of cotton rats with the human respiratory syncytial virus F glycoprotein produced using a baculovirus vector. *Journal of Infectious Diseases* **159**, 255-264.

Wathen, M. W., Brideau, R. J., Thomsen, D. R. and Murphy, B. R. (1989b). Characterization of a novel human respiratory syncytial virus chimaeric FG glycoprotein expressed using a baculovirus vector. *Journal of General Virology* **70**, 2625-2635.

Waxham, M. N., Aronowski, J., Server, A. C., Wolinsky, J. S., Smith, J. A. and Goodman, H. M. (1988). Sequence determination of the mumps virus HN gene. *Virology* **164**, 318-325.

Weir, E. C., Brownstein, D. G., Smith, A. L. and Johnson, E. A. (1988). Respiratory disease and wasting in athymic mice infected with pneumonia virus of mice. *Laboratory Animal Science* **38**, 133-137.

Wertz, G. W., Collins, P. L., Huang, Y., Gruber, C., Levine, S. and Ball, A. L. (1985). Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein. *Proceedings of the National Academy of Sciences, USA* **82**, 4075-4089.

Wertz, G. W., Stott, E. J., Young, K. K. Y., Anderson, K. and Ball, L. A. (1987). Expression of the fusion protein of human respiratory syncytial virus from recombinant vaccinia virus vectors and protection of vaccinated mice. *Journal of Virology* **61**, 293-301.

Wertz, G. W., Krieger, M. and Ball, L. A. (1989). Structure and cell surface maturation of the attachment glycoprotein of human respiratory syncytial virus in a cell line deficient in O-glycosylation. *Journal of Virology* **63**, 4767-4776.

Wild, T. F., Malvoisin, E. and Buckland, R. (1991). Measles virus: both the haemagglutinin and fusion glycoproteins are required for fusion. *Journal of General Virology* **72**, 439-442.

Williams, P. M., Williamson, K. A. and Emerson, S. U. (1988). Monoclonal antibodies to the NS protein of vesicular stomatitis virus inhibit initiation of transcripts in vitro and dissociate leader RNA from mRNA synthesis. *Virology* **167**, 342-348.

Wills, N. M., Gesteland, R. F. and Atkins, J. F. (1991). Evidence that a downstream pseudoknot is required for translational read-through of the Moloney murine

leukemia virus gag stop codon. *Proceedings of the National Academy of Sciences, USA* **88**, 6991-6995.

Wilson, C., Gilmore, R. and Morrison, T. (1990). Aberrant membrane insertion of a cytoplasmic tail deletion mutant of the hemaggluitinin-neuraminidase glycoprotein of Newcastle disease virus. *Molecular and Cellular Biology* **10**, 449-457.

Wilson, W., Malim, M. H., Mellor,J., Kingsman, A. J. and Kingsman, S. M. (1986). Expression strategies of the yeast transposon Ty: a short sequence directs ribosomal frameshifting. *Nucleic Acids Research* **14**, 7001-7015.

Wilson, W., Braddock, M., Adams, S. E., Rathjen, P. D., Kingsman, S. M. and Kingsman, A. J. (1988). HIV expression strategies: Ribosomal frameshifting is directed by a short sequence in both mammalian and yeast systems. *Cell* **55**, 1159-1170.

Xiong, Z. and Lommel, S. A. (1989). The complete nucleotide sequence and genome organization of red clover necrotic mosaic virus RNA-1. *Virology* **171**, 543-554.

Ye, Z., Baylor, N. W. and Wagner, R. R. (1989). Transcription-inhibition and RNA-binding domains of influenza A virus matrix protein mapped with anti-idiotypic antibodies and synthetic peptides. *Journal of Virology* **63**, 3586-3594.

Yu, Q., Davis, P. J., Li, J. and Cavanagh, D. (1992). Cloning and sequencing of the matrix protein (M) gene of turkey rhinotracheitis virus reveal a gene order different from that of respiratory syncytial virus. *Virology* **186**, 426-434.

Yusoff, K., Nesbit, M., McCartney, H., Meulemans, G., Alexander, D. J., Collins, M. S., Emmerson, P. T. and Samson, A. C. R. (1989). Location of neutralizing epitopes on the fusion protein of Newcastle disease virus strain Beaudette C. *Journal of General Virology* **70**, 3105-3109.

Zamora, M. and Samal, S. K. (1992). Sequence analysis of M2 mRNA of bovine respiratory syncytial virus obtained from an F-M2 dicistronic mRNA suggests structural homology with that of human respiratory syncytial virus. *Journal of General Virology* **73**, 737-741.